

Stroud, A. L. (2012) Genetic analysis of RPA single-stranded DNA binding protein in *Haloferax volcanii*. PhD thesis, University of Nottingham.

**Access from the University of Nottingham repository:**

[http://eprints.nottingham.ac.uk/12623/1/Amy\\_Stroud\\_Final\\_corrected\\_version\\_.pdf](http://eprints.nottingham.ac.uk/12623/1/Amy_Stroud_Final_corrected_version_.pdf)

**Copyright and reuse:**

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

- Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners.
- To the extent reasonable and practicable the material made available in Nottingham ePrints has been checked for eligibility before being made available.
- Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
- Quotations or similar reproductions must be sufficiently acknowledged.

Please see our full end user licence at:

[http://eprints.nottingham.ac.uk/end\\_user\\_agreement.pdf](http://eprints.nottingham.ac.uk/end_user_agreement.pdf)

**A note on versions:**

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact [eprints@nottingham.ac.uk](mailto:eprints@nottingham.ac.uk)

**Genetic Analysis of RPA Single-stranded DNA  
Binding Protein in *Haloferax volcanii***

Amy Laura Stroud, BSc Hons.

Thesis submitted to the University of Nottingham  
for the degree of Doctor of Philosophy,  
October 2011

# Contents

<b>Contents.....</b>	<b>i</b>
<b>Abstract .....</b>	<b>vi</b>
<b>Acknowledgements .....</b>	<b>vii</b>
<b>Abbreviations .....</b>	<b>viii</b>
<b>Chapter 1: Introduction .....</b>	<b>1</b>
<b>1.1 Archaea .....</b>	<b>1</b>
1.1.1 The discovery of Archaea.....	1
1.1.2 Characteristics of Archaea .....	2
<b>1.2 The model organism <i>Haloferax volcanii</i> .....</b>	<b>4</b>
<b>1.3 DNA replication.....</b>	<b>9</b>
1.3.1 Bacterial DNA replication .....	10
1.3.2 DNA Replication in Eukaryotes .....	16
1.3.3 Archaeal DNA replication.....	25
<b>1.4 DNA repair .....</b>	<b>33</b>
1.4.1 Nucleotide excision repair.....	34
1.4.2 Global Genome repair.....	34
1.4.3 Transcription-coupled nucleotide excision repair .....	36
1.4.4 Archaeal nucleotide excision repair .....	37
1.4.5 Base excision repair .....	38
1.4.6 Mismatch repair .....	42
1.4.7 Mismatch repair in <i>E. coli</i> .....	43
1.4.8 Mismatch repair in eukaryotes.....	44
1.4.9 Homologous recombination .....	47
1.4.10 Homologous recombination in <i>E. coli</i> .....	51
1.4.11 Homologous recombination in <i>Bacillus subtilis</i> .....	54
1.4.12 Homologous recombination in eukaryotes .....	55
1.4.13 Meiotic recombination .....	60
1.4.14 Homologous recombination in archaea.....	61
1.4.15 Non-homologous end joining .....	62
1.4.16 Non-homologous end joining in eukaryotes.....	64
1.4.17 Non-homologous end joining in <i>E. coli</i> .....	65

1.4.18 Interstrand cross-link repair .....	66
1.4.19 Interstrand cross-link repair in <i>E. coli</i> .....	67
1.4.20 Interstrand cross-link repair in eukaryotes .....	69
<b>1.5 Single stranded DNA binding proteins.....</b>	<b>71</b>
1.5.1 Bacterial single-stranded DNA binding protein (SSB) .....	72
1.5.2 Eukaryotic RPA structure and function .....	74
1.5.3 Yeast RFA .....	78
1.5.4 Archaeal RPA .....	78
<b>Chapter 2: Materials and Methods .....</b>	<b>82</b>
<b>2.1 Materials .....</b>	<b>82</b>
2.1.1 Tables .....	82
2.1.2 <i>Haloferax volcanii</i> media .....	101
2.1.3 <i>Haloferax volcanii</i> buffers and solutions .....	103
2.1.4 <i>Escherichia coli</i> media.....	103
2.1.5 Other buffers and solutions .....	104
<b>2.2 Methods .....</b>	<b>106</b>
2.2.1 Nucleic Acid Manipulation.....	106
<b>Step.....</b>	<b>107</b>
<b>DyNAzyme.....</b>	<b>107</b>
<b>Phusion .....</b>	<b>107</b>
2.2.2 Gel Electrophoresis.....	108
2.2.3 Cloning and gene deletion analysis .....	109
2.2.4 <i>Escherichia coli</i> Microbiology .....	110
2.2.5 <i>Haloferax volcanii</i> microbiology .....	111
<b>Step.....</b>	<b>116</b>
<b>DyNAzyme.....</b>	<b>116</b>
2.2.6 Phenotype analysis.....	116
2.2.7 Protein purification in <i>Haloferax volcanii</i> .....	118
2.2.8 Protein analysis.....	119
<b>Chapter 3: Replication protein A of <i>Haloferax volcanii</i> .....</b>	<b>121</b>
<b>3.1 Phylogenetic analysis of archaeal RPA .....</b>	<b>121</b>
<b>3.2 Expression of the three RPAs .....</b>	<b>126</b>
<b>3.3 Discussion .....</b>	<b>128</b>
<b>Chapter 4: RPA1 .....</b>	<b>129</b>
<b>4.1. Isolation of the <i>rpa1</i> operon.....</b>	<b>129</b>



<b>4.2. Generation of single <i>rpa1</i>, <i>rpa1ap</i> and <i>rpe</i> deletion mutants .....</b>	<b>131</b>
4.2.1. Construction of <i>rpa1</i> deletion mutant.....	134
4.2.2 Construction of <i>rpa1ap</i> deletion mutant.....	138
4.2.3. Construction of <i>rpe</i> deletion mutant.....	143
<b>4.3. Generation of <i>rpa1</i> operon deletion .....</b>	<b>148</b>
4.3.1 Construction of <i>rpa1</i> operon deletion mutant.....	148
<b>4.4. DNA damage assays of deletion mutants.....</b>	<b>151</b>
4.4.1. UV DNA damage assays.....	152
4.4.2. MMC DNA damage assays .....	153
4.4.3. Hydrogen peroxide DNA damage assays .....	154
<b>4.5. Growth competition assay.....</b>	<b>156</b>
<b>4.6. Discussion.....</b>	<b>157</b>
<b>Chapter 5: RPA3 .....</b>	<b>158</b>
<b>5.1. Isolation of the <i>rpa3</i> operon.....</b>	<b>158</b>
<b>5.2 Generation of single <i>rpa3</i> and <i>rpa3ap</i> deletion mutants.....</b>	<b>160</b>
5.2.1 Construction of <i>rpa3</i> deletion mutant.....	160
5.2.2 Construction of <i>rpa3ap</i> deletion mutant.....	164
<b>5.3 Generation of <i>rpa3</i> operon deletion .....</b>	<b>168</b>
5.3.1 Construction of <i>rpa3</i> operon deletion mutant.....	168
<b>5.4 DNA damage assays of deletion mutants.....</b>	<b>171</b>
5.4.1 UV DNA damage assays.....	171
5.4.2 MMC DNA damage assays.....	172
5.4.3 Hydrogen peroxide DNA damage assays .....	173
<b>5.5 Growth competition assay.....</b>	<b>175</b>
<b>5.6 Discussion .....</b>	<b>176</b>
<b>Chapter 6: RPA2 .....</b>	<b>177</b>
<b>6.1 Isolation of the <i>rpa2</i> operon.....</b>	<b>177</b>
<b>6.2 Overexpression of Rpa2.....</b>	<b>179</b>
<b>6.3 Generation of <i>rpa2</i> deletion mutant.....</b>	<b>181</b>
6.3.1 Construction of <i>rpa2</i> deletion mutant.....	181
6.3.2 Construction of <i>trpA</i> marked <i>rpa2</i> deletion mutant.....	184
6.3.3 Construction of <i>rpa2</i> complementation plasmid.....	186
<b>6.4 Integration of <i>rpa2</i> at an ectopic locus .....</b>	<b>188</b>
6.4.1 Construction of the $\Delta rpa2$ <i>hdrB</i> marked plasmid.....	190
6.4.2 Insertion of <i>rpa2</i> at an ectopic locus.....	191

6.5 Insertion of <i>hdrB</i> at the <i>pyrE2</i> locus.....	196
6.5.1 Construction of <i>pyrE2</i> deletion construct marked with <i>hdrB</i> .....	196
6.5.2 Autotrophic conformation of <i>hdrB</i> insertion at the <i>pyrE2</i> locus .....	197
6.6 Discussion .....	199
<b>Chapter 7: Multiple deletion mutants .....</b>	<b>202</b>
7.1. <i>Δrpe Δmre11 Δrad50</i> and <i>Δrpe Δmre11</i> deletion mutants.....	202
7.1.1. Generation of <i>Δrpe Δmre11 Δrad50</i> and <i>Δrpe Δmre11</i> deletion mutants .....	202
7.1.2 DNA damage assays of <i>Δrpe Δmre11 Δrad50</i> and <i>Δrpe Δmre11</i> deletion mutants .....	205
7.2. <i>Δrpa1 Δrpa3ap</i> deletion mutant .....	207
7.2.1. Construction of non- <i>trpA</i> marked <i>Δrpa1</i> deletion mutant.....	207
7.2.2. Attempt to generate double <i>Δrpa1Δrpa3ap</i> deletion mutant.....	208
7.3. <i>Δrpa1ap Δrpa3</i> deletion mutant .....	210
7.3.1. Construction of <i>rpa1ap</i> deletion mutant.....	210
7.3.2. Attempt to generate double <i>Δrpa1apΔrpa3</i> deletion mutant.....	212
7.4. <i>Δrpa1ap Δrpa3ap</i> deletion mutant .....	214
7.4.1. Attempt to generate double <i>Δrpa1ap Δrpa3p</i> deletion mutant.....	214
7.5. <i>Δrpa1</i> operon <i>Δrpa3</i> operon deletion mutants.....	215
7.5.1 Construction <i>rpa1</i> operon complementation plasmid .....	215
7.5.2 Generation of double <i>Δrpa1</i> operon <i>Δrpa3</i> operon deletion mutant.....	216
Growth comparison of the double operon deletion mutant.....	218
7.5.2. DNA damage assays of deletion mutants .....	218
7.6 Discussion .....	221
<b>Chapter 8: Histidine-tagged RPA and associated proteins.....</b>	<b>222</b>
8.1. Construction of histidine-tagged <i>rpa1</i> with <i>rpa1ap</i> .....	224
8.1.1. Construction of histidine-tagged <i>rpa1</i> .....	224
8.1.2. Insertion of untagged <i>rpa1ap</i> downstream of histidine-tagged <i>rpa1</i> .....	226
8.2. Construction of <i>rpa1</i> with histidine-tagged <i>rpa1ap</i> .....	228
8.2.1. Construction of histidine-tagged <i>rpa1ap</i> .....	228
8.1.2. Insertion of untagged <i>rpa1</i> upstream of histidine-tagged <i>rpa1ap</i> .....	230
8.3 Histidine-tag pulldowns of RPA1 and RPA1AP .....	232
RPA1 1C_MP2411 .....	235
8.4. Construction of histidine-tagged <i>rpa3</i> with <i>rpa3ap</i> .....	237
8.4.1. Construction of histidine-tagged <i>rpa3</i> .....	237
8.4.2. Insertion of untagged <i>rpa3ap</i> downstream of histidine-tagged <i>rpa3</i> .....	239

<b>8.5. Construction of <i>rpa3</i> with histidine-tagged <i>rpa3ap</i></b>	<b>240</b>
8.5.1. Construction of histidine-tagged <i>rpa3ap</i>	240
8.5.2. Insertion of untagged <i>rpa3</i> upstream of histidine-tagged <i>rpa3ap</i>	242
<b>8.6 Histidine-tag pulldown of RPA3 and RPA3AP</b>	<b>244</b>
<b>RPA3 1A_MP2511</b>	<b>246</b>
<b>8.7. Construction of histidine-tagged <i>rpa2</i> plasmid</b>	<b>248</b>
8.7.1. Construction of histidine-tagged <i>rpa2</i>	248
<b>8.8 Histidine-tag pulldown of RPA2</b>	<b>249</b>
<b>8.9. Discussion</b>	<b>251</b>
<b>Discussion</b>	<b>252</b>
<b>Future Perspectives</b>	<b>254</b>
<b>Bibliography</b>	<b>256</b>
<b>Appendix</b>	<b>270</b>
<b>Rpa3AP and RPA3AP MP_2511</b>	<b>275</b>
<b>Note in proof</b>	<b>280</b>

## Abstract

Replication protein A (RPA) is a single-stranded DNA-binding protein that is present in all three domains of life. The roles of RPA include stabilising and protecting single-stranded DNA from nuclease degradation during DNA replication and repair. To achieve this, RPA uses an oligosaccharide-binding fold (OB fold) to bind single-stranded DNA.

*Haloferax volcanii* encodes three RPAs – RPA1, RPA2 and RPA3, of which *rpa1* and *rpa3* are in operons with genes encoding associated proteins (APs). The APs belong to the COG3390 group of proteins found in Euryarchaeota and feature an OB fold. Genetic analysis of deletion mutants was employed to determine if all three RPAs are essential for cell viability, and if there is an element of redundancy between RPA1 and RPA3. The hypothesis that the RPAs form a complex with their respective APs, as opposed to a heterotrimeric RPA complex, was also investigated. Furthermore, it was tested whether the RPAs and their respective APs are specific for each other, or whether they are interchangeable.

The genetic analysis showed that RPA2 is essential for cell viability, but that neither RPA1 nor RPA3 are. The *rpa3*, *rpa3ap* and the *rpa3* operon deletion mutants showed sensitivity to DNA damage but only a slight growth defect. By contrast, the *rpa1*, *rpa1ap*, *rpe* and *rpa1* operon mutants did not show any DNA damage sensitivity and an even milder growth defect. The double *rpa1 rpa3* operon deletion was difficult to generate but unexpectedly lacked a significant DNA damage sensitivity and growth defect. The inability to make the double *rpa1 rpa3ap* and *rpa1ap rpa3* deletion mutants suggests that the APs are specific for their respective RPAs.

Biochemical analysis involving histidine-tagged RPAs and APs was used to confirm the conclusions of the genetic analysis. The RPAs did not interact with each other, but instead co-purified with their respective APs. This finding reiterates that the RPAs do not form a heterotrimeric complex, as seen in eukaryotes, but instead form a novel complex with their respective APs.

## **Acknowledgements**

I would like to say thank you to my supervisor Thorsten Allers, for all his guidance and support, and to Liz Sockett for her wisdom. Many thanks to Susan Liddell and all the members of C5, past and present, for their advice and help, particularly Michelle and Kayleigh.

I am eternally grateful to my Mother and Father, for their endless love and support, I could not of succeeded without them. A special thank you to Nanny, Pappy and Pete, for being there and spurring me on. Finally, I will always be deeply grateful to Charlie and Gwenny, for making the impossible seem possible

## Abbreviations

5-FOA .....	5-fluoroorotic acid
ACS .....	autonomously replicating sequence consensus sequence
ARS .....	autonomously replicating sequence
BER .....	base excision repair
BLAST .....	basic local alignment search tool
bp .....	base pair
CDK .....	cyclin-dependent kinases
COG .....	cluster of orthologous groups
DDK .....	Dbf4-dependent kinases
DNA .....	deoxyribose nucleic acid
dNTP .....	deoxynucleotide
DSBR .....	double-stranded break repair
dsDNA .....	double-stranded deoxyribose nucleic acid
DUE .....	duplex unwinding element
EDTA .....	ethylenediaminetetraacetic acid
FA .....	Fanconi anemia
H <sub>2</sub> O <sub>2</sub> .....	hydrogen peroxide
HTH .....	helix-turn-helix
HJ .....	Holliday junction
HR .....	homologous recombination
HTH .....	helix-turn-helix
Hv-Ca .....	<i>Haloferax volcanii</i> casamino acids media
Hv-YPC .....	<i>Haloferax volcanii</i> yeast extract, peptone and casamino acids media
ICL .....	interstrand crosslink
kb .....	kilobase
MCM .....	mini-chromosome maintenance
MMC .....	mitomycin C
MMR .....	mismatch repair
NER .....	nucleotide excision repair
NHEJ .....	non-homologous end joining
ORB .....	origin recognition box

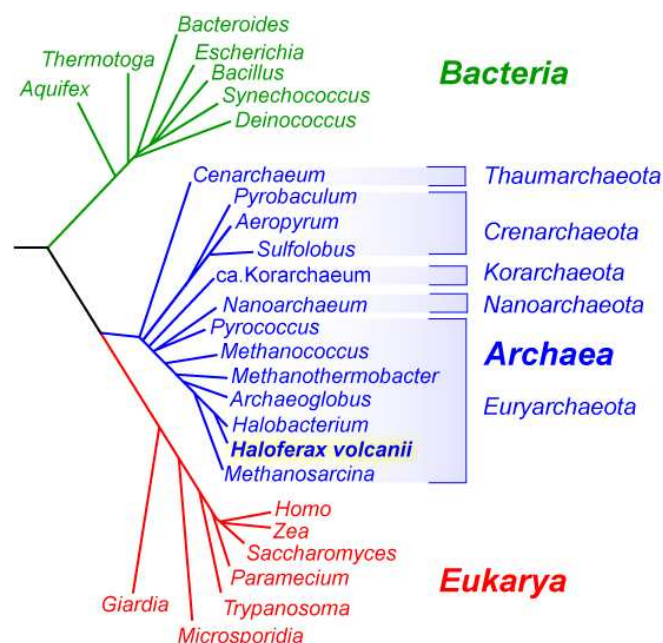
ORC .....	origin recognition complex
PCNA .....	proliferating cell nuclear antigen
PCR .....	polymerase chain reaction
RPA .....	replication protein A
RNA .....	ribonucleic acid
rRNA .....	ribosomal deoxyribose nucleic acid
RT PCR .....	reverse transcriptase polymerase chain reaction
SDSA .....	synthesis-dependent strand annealing
SSA .....	single-strand annealing
SSB .....	single-strand DNA binding protein
ssDNA .....	single-stranded deoxyribose nucleic acid
TIGR .....	The Institute for Genomic Research
TFIIH .....	transcription factor II H
thy .....	thymidine
trp.....	tryptophan
ura .....	uracil
UV.....	ultraviolet light
v/v .....	volume per volume
w/v .....	weight per volume
WT .....	wild-type
X-gal .....	X-galactosidase
XP .....	xeroderma pigmentosum

# Chapter 1: Introduction

## 1.1 Archaea

### 1.1.1 The discovery of Archaea

The development of nucleic acid sequencing techniques in the 20<sup>th</sup> century led to the identification of Archaeabacteria by Carl Woese and George Fox in 1977. Sequencing of the small subunit of the ribosomal RNA, which has a universal presence in living organisms, allowed a sophisticated phylogenetic analysis of a class of distinct anaerobic ‘bacteria’ that did not rely on cell morphology, physiology or pathogeneticity. This revealed that this class of anaerobic ‘bacteria’ was both equally related to bacteria and eukaryotes. The anaerobic ‘bacteria’ were consequently renamed Archaeabacteria, establishing that the living world is tripartite not bipartite (Woese and Fox 1977). The isolation and characterization of more Archaeabacteria revealed that they are related more closely to eukaryotes than bacteria, and consequently Archaeabacteria were renamed Archaea. Furthermore, it was put forward that Eukarya, Archaea and Bacteria would be known as the three domains of life (Woese, Kandler *et al.* 1990)



**Figure 1.1 Tree of life**

Phylogenetic tree of the three domains of life based on rRNA sequences with selected phyla and genre shown. Adapted from (Allers and Mevarech 2005) with modifications from (Brochier-Armanet, Forterre *et al.* 2011).



The phylogenetic analysis of ribosomal RNA (rRNA) sequencing showed that archaea cluster into several phyla, as shown in Figure 1.1, the Korarchaeota and Thaumarchaeota are the most recent to be established (Brochier-Armanet, Forterre *et al.* 2011). Significant biological differences can be seen between the phyla, for example the histone proteins, extremely similar to those in eukaryotes, are present in Euryarchaeota, Thaumarchaeota and a single Crenarchaeota *Thermoproteales*. Conversely, only Thaumarchaeota possess the eukaryotic-like DNA Topoisomerase IB protein. (Brochier-Armanet, Forterre *et al.* 2011).

### **1.1.2 Characteristics of Archaea**

Many archaea possess the ability to live in harsh and unusual environments with extreme pH, salinity or temperature, which would otherwise be barren of life. For example, halophilic archaea grow extremely well in salt lakes such as the Dead Sea. Despite the extreme environments archaea inhabit, their presence is significantly widespread, and mesophilic archaea account for 20% of the total biomass (DeLong and Pace 2001).

Methanogens, the first archaea to be discovered, make up five orders of the *Euryarchaeota*: *Methanococcales*, *Methanosarcinales*, *Methanobacteriales*, *Methanomicrobiales* and *Methanopyrales* (Leigh, Albers *et al.* 2011). Methanogens are able to generate methane as a catabolic end product using H<sub>2</sub> and CO<sub>2</sub> as a substrate, and inhabit environments such as rice paddies, sewage sludge digesters and hydrothermal vents (Wolfe 1996; Leigh, Albers *et al.* 2011). The *Archaeaoglobales*, *Pyrococcales* and *Methanococcales* orders feature hyperthermophiles, which in the case of *Archaeaoglobus* are sulphur-reducing as well. Hyperthermophiles grow between 50 and 100 °C, typically >80 °C, with *Pyrococcus* growing optimally between 95 and 105 °C (Leigh, Albers *et al.* 2011). The order *Sulfolubales* is characterized by aerobic thermoacidophiles, which inhabit acidic hot springs and mudholes, typically growing at 70-80 °C and at a pH value of 2-3 (Leigh, Albers *et al.* 2011). Halophilic archaea, including *Halobacterium* and *Haloferax volcanii*, inhabit extreme saline environments including solar salterns and natural salt lakes. Halophilic archaea are adapted to high salt conditions maintaining the osmotic balance by having a hypersaline cytoplasm, which is achieved through a ‘salt-in’ strategy using potassium over sodium as it attracts less water (Christian and Waltho 1962) (Oren

2008). Halophilic proteins adapt to this hypersaline environment by incorporating hydrophilic residues on their surfaces (Wright, Banks *et al.* 2002) and a high percentage of acidic residues that provide negative charges to coordinate a network of hydrated cations that allow a protein to remain soluble (Lanyi 1974). An extra acidic domain in proteins, which is not present in mesohalic proteins, have also been identified that aid correct protein folding in high salt conditions, for example the ferredoxin of *Halobacterium salinarum* (Marg, Schweimer *et al.* 2005). Halophiles grow aerobically and optimally at 40-50°C (Robinson, Pyzyrna *et al.* 2005) and many genomes of haloarchaea have been shown to possess multiple isoforms of genes, for example *H. volcanii* possesses sixteen DNA replication initiator *orc1/cdc6I* genes (Norais, Hawkins *et al.* 2007).

As will be discussed in greater detail later, the archaeal proteins involved in information processing systems, such as DNA replication and repair, show great similarity to their eukaryotic counterpart. For example, the core components of the archaeal RNA polymerase II are closely related to the eukaryotic homolog (Huet, Schnabel *et al.* 1983). The presence and the tertiary structure of the archaeal histones is significantly similar to the eukaryotic histones (White and Bell 2002), and although the circular genome structure of archaea is similar to that of bacteria (Grabowski and Kelman 2003), the DNA replication process is more similar to that seen in eukaryotes. Coupled with a central metabolism where energy conversion and biosynthesis processes are closely related bacterial counterparts (Rivera, Jain *et al.* 1998), archaea are seen as a chimera of eukaryotic and bacterial elements. The observation of eukaryotic traits present throughout archaea provided grounds for the argument that eukaryotes originated from an archaeal lineage. Conversely, the lack of a single archaeon possessing all of the eukaryotic traits and the failure of large-scale genomic analyses to identify a consensus on the evolutionary relationship between archaea and eukaryotes has sparked debate. This has led to the alternative hypothesis that the eukaryotic features observed in archaea were also present in a common ancestor, which was more complex than its present day archaeal descendants (Brochier-Armanet, Forterre *et al.* 2011).

## 1.2 The model organism *Haloferax volcanii*

The choice of a model organism is extremely important, not only must it be easy to grow with a short generation time, but the amenability of a model organism to experimental manipulation is also critical (Leigh, Albers *et al.* 2011). All three of these requirements can be found in *Haloferax volcanii*.

The haloarchaeon *H. volcanii* was originally isolated from the Dead Sea (Mullakhanbhai and Larsen 1975) *H. volcanii* has a disc-shaped morphology, has a red pigmentation due to the presence of carotenoids, and as discussed earlier is biased in acidic amino acids compared to basic amino acids to allow adaption to high salt environments (Leigh, Albers *et al.* 2011) (Lanyi 1974). The GC content of *H. volcanii* is ~66% (Hartman, Norais *et al.* 2010), and the ploidy of *H. volcanii* ranges from 10 copies per cell at stationary phase to 18 copies per cell during exponential phase (Breuert, Allers *et al.* 2006). Instead of a rigid cell wall *H. volcanii* possesses a single layer of glycoproteins called the Surface layer (S-layer). *H. volcanii* grows optimally at 45°C (Mullakhanbhai and Larsen 1975) with a generation time of 2.8 hours, and can be grown in liquid or solid media containing 2.5 M NaCl. The genomic structure of the *H. volcanii* genome consists of a main circular chromosome (2.848 Mb) and four smaller replicons pHV4 (636 kb), pHV3 (438 kb), pHV1 (85.1 kb) and pHV2 (6.35 kb) (Charlebois, Schalkwyk *et al.* 1991) (Hartman, Norais *et al.* 2010). A wide range of genetic tools for *H. volcanii* has been developed, shown in Table 1.1, many of which revolve around introducing DNA into *H. volcanii* cells (Leigh, Albers *et al.* 2011).

<b>Genetic Tools</b>	<b><i>Hfx. volcanii</i></b>
<b>Synthetic media</b>	Yes (Mevarech and Werczberger 1985)
<b>DNA delivery</b>	PEG-mediated transformation (Cline, Lam <i>et al.</i> 1989)
<b>Restriction barrier</b>	Severe restriction of G <sup>me</sup> ATC DNA, eliminated in $\Delta mrr$ mutant (Holmes and Dyall-Smith 1991; Allers, Barak <i>et al.</i> 2010)
<b>Replicative shuttle vectors</b>	Based on pHK2, pHV2 and pHV1/4 origins (Lam and Doolittle 1989), (Holmes, Pfeifer <i>et al.</i> 1994), (Allers, Ngo <i>et al.</i> 2004), (Allers and Mevarech 2005), (Norais, Hawkins <i>et al.</i> 2007)
<b>Positive selection</b>	Mevinolin, novobiocin, uracil ( <i>pyrE2</i> ), leucine ( <i>leuB</i> ), thymidine ( <i>hdrB</i> ), tryptophan ( <i>trpA</i> ) (Lam and Doolittle 1989), (Holmes and Dyall-Smith 1991), (Bitan-Banin, Ortenberg <i>et al.</i> 2003), (Allers, Ngo <i>et al.</i> 2004). Also histidine ( <i>hisC</i> ) and methionine ( <i>metX</i> ) (M. Mevarech, personal communication)
<b>Counter-selection</b>	5-FOA ( <i>pyrE2</i> ) (Bitan-Banin, Ortenberg <i>et al.</i> 2003; Allers, Ngo <i>et al.</i> 2004)
<b>Random mutagenesis</b>	Using ethyl methanesulphonate (EMS) (Mevarech and Werczberger 1985)
<b>Negative enrichment</b>	Using 5-bromo-2'-deoxyuridine (BrdU) (Soppa and Oesterhelt 1989) (Wanner and Soppa 1999)
<b>Markerless gene knockout or replacement</b>	Using <i>pyrE2</i> (Bitan-Banin, Ortenberg <i>et al.</i> 2003; Allers, Ngo <i>et al.</i> 2004). Gateway system available (El Yacoubi, Phillips <i>et al.</i> 2009)
<b>Ectopic integration</b>	At <i>pyrE2</i> (T. Allers, unpublished), see Chapter 6.
<b>Natural genetic exchange</b>	Involves cell-cell contact (Rosenshine and Mevarech 1989)
<b>Reporter genes</b>	$\beta$ -galactosidase ( <i>bgaH</i> ) and GFP (Holmes and Dyall-Smith 2000; Reuter and Maupin-Furlow 2004)
<b>Regulated gene expression</b>	Tryptophan-inducible p. <i>tnaA</i> promoter (Large, Stämme <i>et al.</i> 2007)
<b>Protein overexpression</b>	<i>pitA<sub>Nph</sub></i> gene replacement strain H1209 and tryptophan-inducible expression plasmid pTA963, for His-tagged proteins (Allers, Barak <i>et al.</i> 2010)

**Table 1.1 Genetic tools for *H. volcanii***

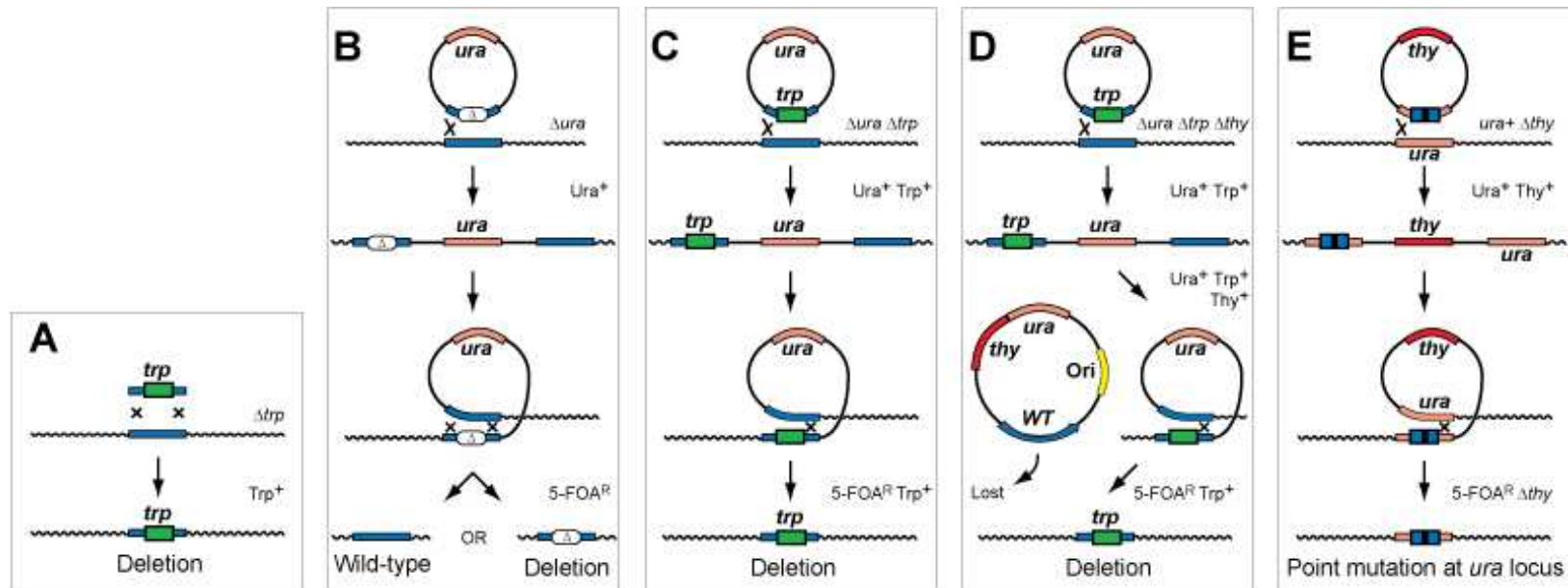
Genetic tools that can be used to genetically manipulate *H. volcanii*. Taken from (Leigh, Albers *et al.* 2011).

DNA can be transformed into *H. volcanii* by using the polyethylene glycol-mediated transformation (Cline, Schalkwyk *et al.* 1989). EDTA is used to remove the S-layer through chelation of magnesium and DNA is then introduced into resulting

spheroplasts using PEG 600. The cells are then allowed to recover in rich medium and plated on selective media. The curing of the pHV2 replicon in the wild type laboratory strain DS70 (Wendoloski, Ferrer *et al.* 2001), has allowed the pHV2 shuttle plasmid to be used for genetic analysis by transformation into *H. volcanii* (Wendoloski, Ferrer *et al.* 2001) (Lam and Doolittle 1989). Other shuttle vectors based on pHK2 and PHV1/4 origins have been developed as well.

Antibiotic markers include novobiocin (Holmes and Dyall-Smith 1991), which inhibits the DNA gyrase, and mevinolin (simvastatin), a HMG-CoA reductase inhibitor (Lam and Doolittle 1992). However both antibiotic resistance markers have homology to essential genes, which is a problem when using homologous recombination to integrate a DNA fragment containing either of the markers onto the chromosome. This has been circumvented by developing auxotrophic mutants, to allow the use of autotrophic markers (Allers and Ngo 2003) to select for the presence of DNA transformed into *H. volcanii*. These include uracil (Bitan-Banin, Ortenberg *et al.* 2003), tryptophan, leucine and thymidine (Allers, Ngo *et al.* 2004), the former of which can be used in creating a gene knockout or replacement by counterselection using 5-fluoroorotic acid (5-FOA). Transformation of a deletion construct is selected for using uracil prototrophy and the subsequent loss of the construct and generation of a deletion is selected for using 5-FOA, which in a *ura*<sup>+</sup> strain is converted to toxic 5-fluorouracil. Additional autotrophic markers histidine and methionine have recently been developed (Leigh, Albers *et al.* 2011). The inducible tryptophanase promoter *p.tnaA* allows tight regulation of expression, which is controlled by the levels of tryptophan. This can be used to overexpress a particular gene or to lower the expression of an essential gene (Large, Stamme *et al.* 2007) (Leigh, Albers *et al.* 2011). Two colorimetric reporter genes,  $\beta$ -galactosidase and green fluorescent protein (GFP) can be used to observe the expression of a gene (Leigh, Albers *et al.* 2011).  $\beta$ -galactosidase generates a blue colour from X-gal (Holmes and Dyall-Smith 2000) and GFP fluoresces in *H. volcanii* cells (Reuter and Maupin-Furlow 2004). Due to the negative surface charge of *H. volcanii* proteins, which is critical for their solubility and correct protein-folding, a system has been developed for the conditional overexpression of *H. volcanii* proteins. This includes a plasmid containing a tryptophan-inducible promoter and a hexahistidine tag, as well as the development of

a strain where proteins naturally containing histidine clusters have been modified (Allers, Barak *et al.* 2010). The complete sequencing of the *H. volcanii* genome (Hartman, Norais *et al.*) by The Institute for Genomic Research (TIGR) allows precise gene modifications to be made using the genetic tools described (Allers and Mevarech 2005) (Leigh, Albers *et al.* 2011).



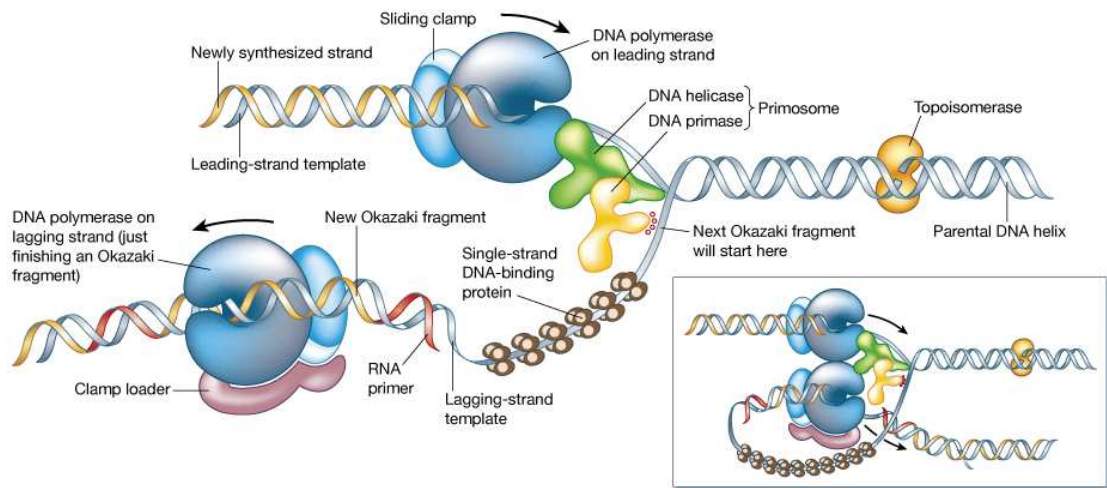
**Figure 1.2 Gene knockout methods in *H. volcanii***

(A) Gene replacement with a selectable marker via homologous recombination between flanking regions present on linear DNA and chromosomal target. (B) Pop-in/pop-out deletion method, where deletion construct integrates on the chromosome and is selected for using uracil prototrophy. Intramolecular recombinants that have lost the plasmid (pop-out) are counter-selected using 5-FOA. (C) Pop-in/pop-out method using a selectable *trp* marker to replace the gene, allowing direct selection at pop-out stage for deletion over wild-type. (D) For genes that are required for cell viability, gene function is complemented *in trans* on a shuttle vector. Both loss of the shuttle vector and gene deletion are achieved by counterselection with 5-FOA. (E) Ectopic integration of a point mutated experimental gene at the *ura* locus. Pop-in is selected for by transformation to thymidine prototrophy, counterselection using 5-FOA ensure *ura* gene is replaced point mutated experimental gene. Taken from (Leigh, Albers *et al.* 2011).

### 1.3 DNA replication

The purpose of DNA replication is to duplicate the genome to allow the proliferation of life to continue. For these reasons DNA replication is an essential and an exact process and one that is tightly regulated. In order for the genome to be entirely and equally duplicated before each cell division, the replicon model provides an accurate and regulated mechanism by which this can occur. The replicon model involves *trans*-acting factors called initiators binding specifically to *cis*-acting sequences called replicators (Jacob and Brenner 1963). By initiating DNA replication at specific sites, known as DNA replication origins, the cell can ensure that only one round of replication from an origin occurs per cycle. This is essential in eukaryotes where multiple origins and multiple origin recognition (ORC) proteins exist, and to ensure replication of large chromosomes. The binding of initiators to the origin promotes DNA replication by recruiting core members of the DNA replication machinery that form the replisome, as shown in Figure 1.3. The replisome opens up the DNA helix by recruiting the replicative helicase in a complex with accessory factors. The resulting ssDNA is stabilized by the single-stranded DNA-binding protein, the resulting helical torsion generated by the helicase activity is removed by a topoisomerase. Unwinding of the double-stranded DNA (dsDNA) allows the clamp loader to load the sliding clamp, a processivity factor that tethers the DNA polymerase to DNA. However the DNA polymerase is unable to initiate *de novo* synthesis and for this reason a RNA polymerase known as a primase synthesises short stretches of RNA *de novo* known as RNA primers. The DNA polymerase is then able to elongate the RNA primers, initiating DNA synthesis. (Sclafani and Holzen 2007). Two DNA polymerases that function on complementary DNA strands are recruited to the replisome to allow the DNA synthesis from the leading and the lagging strand. Due to the 5'-3' direction of the DNA polymerase and the antiparallel structure of duplex DNA, the synthesis of the leading (5'-3') DNA strand is continuous but synthesis of the lagging (3'-5') strand is discontinuous. This leads to the formation of Okazaki DNA fragments on the lagging strand for every time the replication fork is re-loaded (Kornberg and Baker 1992).





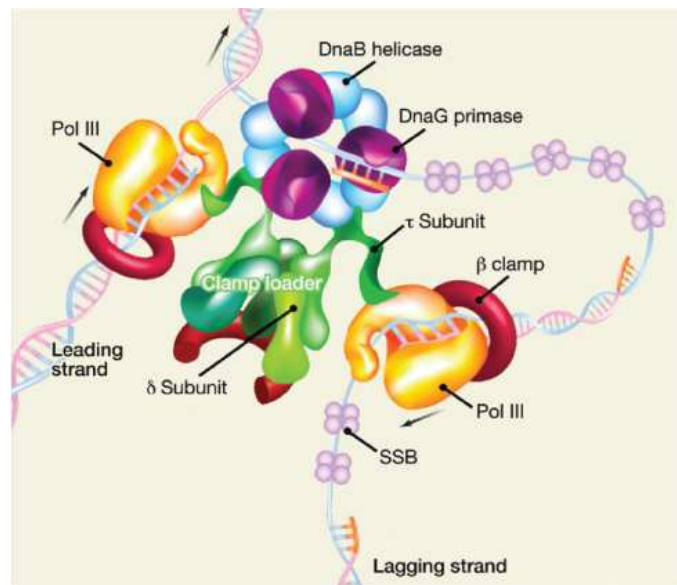
**Figure 1.3 The DNA replisome**

Taken from (Alberts 2003). The main diagram above shows the DNA helicase unwinding the duplex DNA providing a ssDNA template for the DNA primase to synthesise RNA primers. The single-stranded DNA binding protein binds the ssDNA to provide protection from nuclease degradation and from DNA secondary structures forming, which are then extended by the DNA polymerase. The discontinuous synthesis by the DNA polymerase on the lagging strand leads to the formation of Okazaki DNA fragments. The DNA polymerase is tethered to DNA by the sliding clamp that is loaded onto DNA by the clamp loader. The topoisomerase removes any supercoiling generated by the translocation of the DNA helicase.

### **1.3.1 Bacterial DNA replication**

The origin of replication in bacteria is an AT rich region designated *oriC* and characteristically there is one origin for every bacterial chromosome, which are most commonly circular (Kornberg and Baker 1992). The bidirectional replication of the chromosome in *Escherichia coli* is initiated at the *oriC*, a  $\approx 250$  bp region that contains a 9 bp sequence repeated multiple times that makes up the DnaA box, the number of DnaA boxes present in *oriC* differs from species to species (Fuller, Funnell *et al.* 1984; Kornberg and Baker 1992; Robinson and Bell 2005). The DnaA box in *E. coli* is located between two promoters, and is referred to as the DnaA assembly region (DAR) (Speck, Weigel *et al.* 1999) (Mott, Erzberger *et al.* 2008). The *E. coli* initiator protein DnaA is a member of the AAA+ family of ATPases, which possesses a nucleotide-binding domain that facilitates the binding and hydrolysis of ATP (Speck, Weigel *et al.* 1999). Multiple DnaA proteins bind the DAR, forming a large nucleoprotein complex, predominantly on the upper strand of the origin and in the presence of ATP. Resulting in DnaA forming a right-handed spiral filament,

contacting the DNA via the phosphodiester backbone. This allows DnaA to then stretch the contacted DNA strand and thus ‘melt’ the AT rich regions (L,M and R) of the DNA-unwinding element (DUE) neighbouring the DnaA box(s), thereby forming the “open complex” (Fuller, Funnell *et al.* 1984; Bramhill and Kornberg 1988; Schaper and Messer 1995; Speck and Messer 2001; Duderstadt, Chuang *et al.* 2011). The exact mechanism of how DnaA stretches the DUE element is still unclear; Ozaki *et al.* have put two hypotheses forward. The first comprises of DnaA binding the DAR and forming a DnaA nucleoprotein, which then binds the DUE through integration host factor (IHF)-dependent DNA bending. An alternative open complex mechanism involves interaction between DnaA nucleoprotein formed on the DUE and the DnaA nucleoprotein formed on the DAR (Ozaki and Katayama 2011).



**Figure 1.4 The *E. coli* replisome**

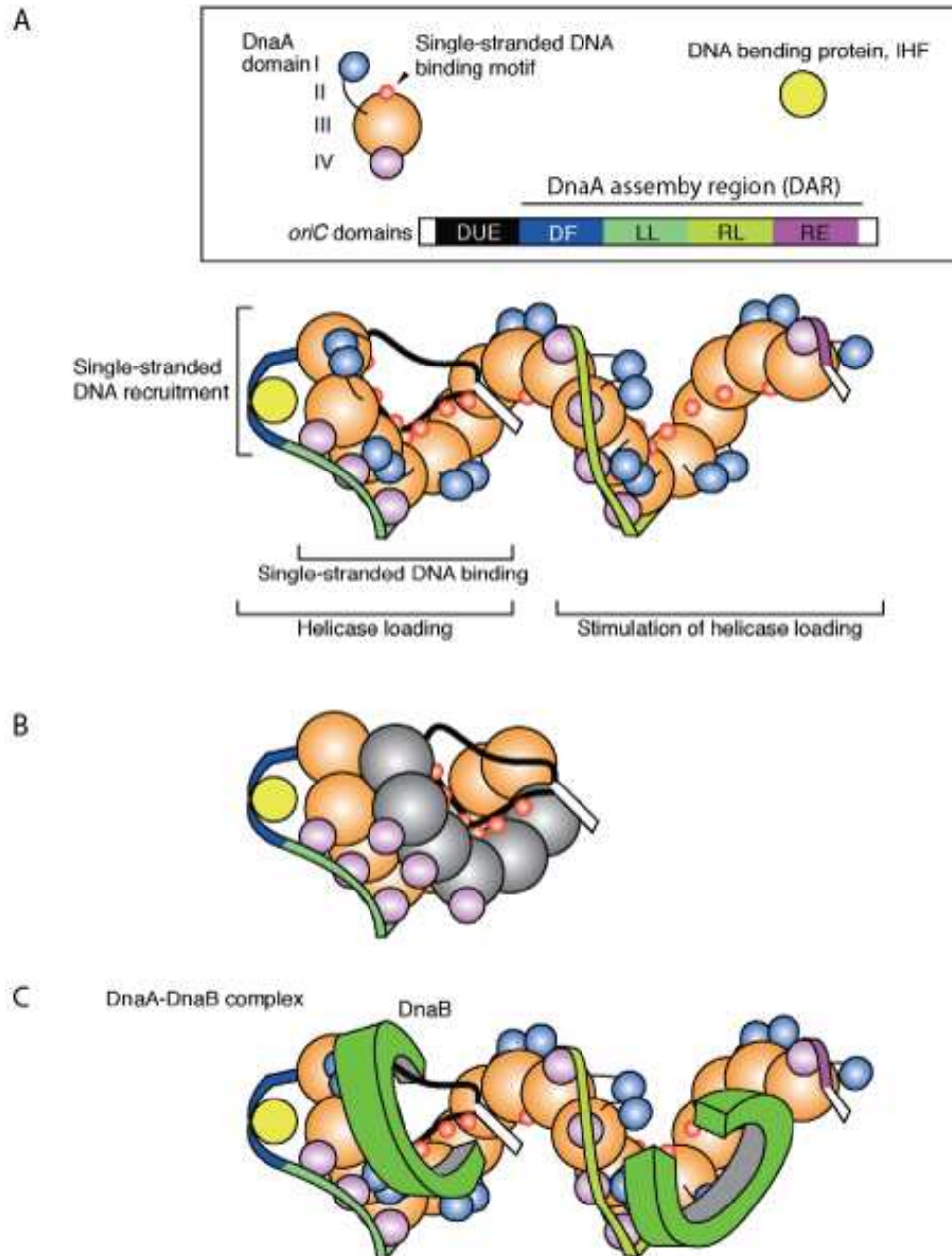
The structure of the *E. coli* replisome, adapted from (Yao and O'Donnell 2010). See text for further details.

The dsDNA is unwound by the replicative helicase DnaB, also a member of the AAA+ ATPase family, and is loaded through DnaC:DnaA interaction and direct DnaB:DnaA interaction (Fang, Davey *et al.* 1999) (Kornberg and Baker 1992). DnaC, a close paralog of DnaA, uses ATP-activated DnaA as a docking site to facilitate the DnaC ssDNA binding and formation of DnaC helical filaments, in an ATP-dependent manner. The DnaA nucleofilament forms off to one side of the DUE on the upper

strand, where DnaC directly interacts with DnaA. The DnaA:DnaC interaction coordinates the loading of DnaB through DnaA N-terminus binding the N-terminus of DnaB and DnaC N-terminus binding of DnaB C-terminus. Thus loading DnaB onto the upper strand of the melted origin. Direct DnaA:DnaC interaction facilitates the loading of the second DnaB helicase onto the lower strand of the melted origin (Mott, Erzberger *et al.* 2008). A high affinity of ATP-DnaC for ssDNA results in extremely strong binding of DnaB hexamers to the origin, suppressing the helicase activity of DnaB (Fang, Davey *et al.* 1999). Hydrolysis of ATP-DnaC to ADP-DnaC releases DnaB allowing both the replicative helicase to become functional and the formation of the prepriming complex (Fang, Davey *et al.* 1999; Davey, Fang *et al.* 2002). In *Bacillus subtilis* there are similarities in the helicase loading process despite differences in the nomenclature; the DNA helicase DnaC is loaded by the dual DnaI-DnaB loader (Grainger, Machon *et al.* 2010).

The ssDNA generated by the action of DnaB helicase is bound by the ssDNA-binding protein (SSB) to prevent the formation of secondary structures and degradation by nucleases. The *E. coli* SSB is homotetramer, made up of 20 kDa subunits, that binds ssDNA using an oligosaccharide binding fold (OB-fold) and has an acidic C-terminal tail that is required for protein-protein interactions (Shereda, Kozlov *et al.* 2008). The resulting ssDNA from the action of DnaB provides a template for the primase DnaG for the *de novo* synthesis of RNA primers, initiation of the primase requires DnaG-DnaB protein interaction and results in primase synthesis in the opposite motion to the DnaB hexamers (Fang, Davey *et al.* 1999). DnaG is made up of three domains: the N-terminal zinc binding domain (ZBD), the central RNA polymerase domain (RPD) and the C-terminal helicase-binding domain (HBD) (Kornberg and Baker 1992; Larson, Griep *et al.* 2010) (Bailey, Eliason *et al.* 2007). DNA primase initiates *de novo* synthesis on defined trinucleotides, *E. coli* DnaG initiates primers synthesis on 5'-d(CTG) whilst *Aquifex aeolicus* DnaG catalyzes primer synthesis on 5'-d(CCC). The ZBD domain of DNA primase is believed to be responsible for the template initiation specificity (Larson, Griep *et al.* 2010). RNA primers synthesised by DnaG are typically 11 nucleotides long, but can range from 10 to 60 nucleotides in length, beginning with pppAG at the 5' end (Frick and Richardson 2001)

The DNA polymerase III holoenzyme is made up of two core DNA polymerases that are linked by a pentameric clamp loader ( $\tau_2\gamma_3$ ), where the  $\tau$  subunits interact with the two polymerase III cores. (Kelman and O'Donnell 1995) (Kornberg and Baker 1992; Onrust, Finkelstein *et al.* 1995; Yuzhakov, Turner *et al.* 1996). The polymerase III core is made up of DNA polymerase  $\alpha$  and  $\epsilon$ , 3'-5' exonuclease and  $\theta$  subunits that are responsible for DNA polymerization and proof-reading. Each core is loaded by a  $\beta$  clamp (Kelman and O'Donnell 1995), a ring-like homodimer of  $\beta$  subunits, encircles DNA and has the ability to travel freely in a sliding motion (Kuriyan and O'Donnell 1993; Kelman and O'Donnell 1995; Yuzhakov, Turner *et al.* 1996). Interaction between the  $\beta$  clamp and the polymerase III core complex stabilizes the polymerase III core complex and the template DNA increasing the processivity of DNA synthesis (Katayama and Sekimizu 1999; Turner, Hingorani *et al.* 1999). The  $\beta$  sliding clamp is loaded by the  $\gamma$  complex, which is made up of  $\gamma$ ,  $\delta$ ,  $\delta'$ ,  $\chi$  and  $\varphi$  subunits. The  $\gamma$  clamp loader binds with a high affinity to the 3' end of the RNA primer when bound to ssDNA, and to the  $\beta$  clamp resulting in the opening of the  $\beta$  clamp and loading onto the primed ssDNA via ATP hydrolysis (Yuzhakov, Turner *et al.* 1996; Hingorani and O'Donnell 1998; Turner, Hingorani *et al.* 1999). The  $\gamma$  clamp loader then disassociates and the polymerase III core complex binds the  $\beta$  clamp, is loaded and DNA replication of the template DNA begins (Kornberg and Baker 1992; Turner, Hingorani *et al.* 1999).



**Figure 1.5 Unwinding of *oriC* and helicase loading**

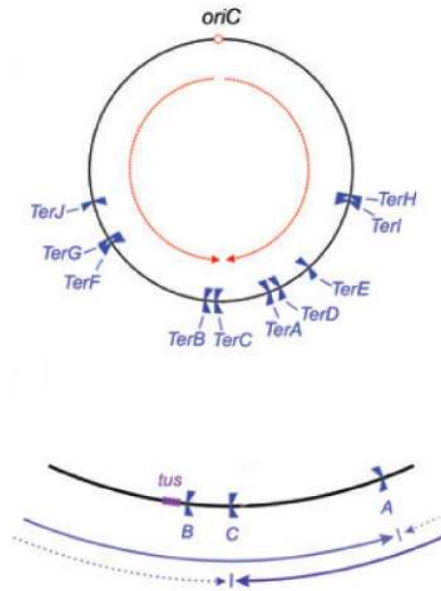
A) The open complex model. DnaA domains (I-IV) and the *oriC* domains (DUE, DF, LL, RL and RE) are shown. The DnaA forms a multimer on the DAR and proceeds to bind the single-stranded DUE (ssDUE), generated by IHF-dependent DNA bending in addition to interaction of DnaA molecules bound to the DF and LL regions. B) Alternative open complex model. DnaA multimers form on the ssDUE (coloured in grey) and the DAR regions, interaction between the two multimers results in the 'stretching' of the DUE. The DF region and DnaA domains I and II are omitted for simplicity. C) Model for helicase loading. The DnaB binding domain on DnaA resides in domain I, which is connected to the AAA<sup>+</sup> domain (domain III) via a flexible linker (domain II), facilitating a flexible interaction between DnaA and DnaB as well as accommodating conformational changes in the *oriC*-DnaA-DnaB-DnaC complex (DnaC has been omitted for simplicity). Adapted from (Ozaki and Katayama 2011).

As discussed earlier DNA synthesis is continuous on the 5'-3' leading strand and discontinuous on the 3'-5' lagging strand (Kornberg and Baker 1992). This contact between DnaB and the leading polymerase core allows the lagging strand polymerase core to hop from one  $\beta$  clamp to another after completion of each Okazaki fragment (Yuzhakov, Turner *et al.* 1996). The  $\beta$  clamps are cycled between the Okazaki fragments by the  $\gamma$  complex to ensure tethering of the polymerase core to each Okazaki fragment, and thus discontinuous replication of the lagging strand (Stukenberg, Turner *et al.* 1994).

Two other DNA polymerases are also involved in DNA replication in *E. coli*, DNA polymerase I and II. DNA polymerase I was the first DNA polymerase discovered in *E. coli* and possesses 3'-5' proof reading exonuclease and 5'-3' exonuclease activities in addition to DNA-dependent 5'-3' DNA polymerization (Kornberg and Baker 1992; Xie and Sayers 2011). The main function of DNA polymerase I is the processing of Okazaki fragments, by removing the RNA primer and replacing with DNA (Nagata, Mashimo *et al.* 2002). DNA polymerase II functions as both a DNA polymerase and a 3'-5' exonuclease but lacks 5'-3' exonuclease activity and is involved in DNA repair, as is DNA polymerase I (Kornberg and Baker 1992).

Termination of DNA replication in *E. coli* occurs at ten 23 bp *Ter* (terminus) sites (*TerA-J*) also known as fork trap sites, which are approximately opposite the origin *oriC* on a circular map shown in figure 1.6. The *Ter* sites are oppositely orientated and bound by the 36 kDa terminator protein Tus. Each *Ter* is only partially effective, the combined effect of each *Ter* site results in the termination of DNA. (Kornberg and Baker 1992) (Duggin, Wake *et al.* 2008) (Mulcair, Schaeffer *et al.* 2006). The first of the two DNA replication forks encounter the permissive face of Tus bound to each of the first series of *Ter* sites. At the forefront of the replisome, the DNA helicase (DnaB), displaces the bound Tus and the replication fork passes through to a second set of *Ter* sites bound by Tus. At this stage the replication fork encounters the nonpermissive face of the Tus resulting in the trapping of the replication fork between oppositely oriented *Ter* sites. A direct interaction between Tus and DnaB prevents the unwinding of *Ter* in a polar manner (Mulcair, Schaeffer *et al.* 2006). A similar

mechanism is used in *Bacillus subtilis* but shows a lack of homology to both the *E. coli* termination site and the termination protein (Duggin, Wake *et al.* 2008)



**Figure 1.6 Termination sites of *Escherichia coli***

Path of replication forks from *oriC* shown by the red lines, the shape of the *Ter* sites denotes their orientation. The enlargement of the termination region shows the terminator protein Tus bound and how opposing *Ter* sites function as a trap. Adapted from (Duggin, Wake *et al.* 2008).

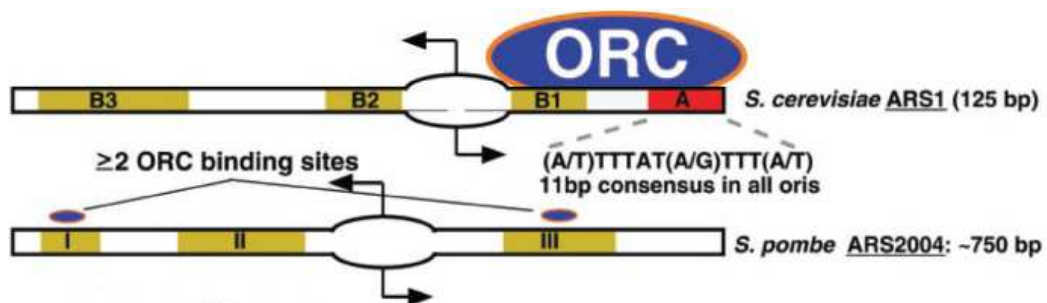
### **1.3.2 DNA Replication in Eukaryotes**

DNA replication in eukaryotes occurs from multiple origins, tight regulation prevents each individual origin from being activated more than once in a cell cycle (DePamphilis, Blow *et al.* 2006). This is an essential requirement due to the vast genome size in comparison to that of bacteria, and a necessity due to multiple chromosomes and multiple origins of replication. This allows larger genomes to be replicated in a shorter time but requires precise coordination in order for mitosis to proceed correctly (Remus and Diffley 2009). DNA replication begins in the G1 phase of the cell-cycle where the origin is recognised and bound by a origin recognition complex (ORC). This transition of the origin from a post-replicative state to a pre-



replication state occurs in the M/G1 phase of the cell cycle and includes the licensing of the origin.

*Saccharomyces cerevisiae* DNA origins were first generically characterized in a study examining the transformation efficiency of plasmids containing DNA fragments and a selectable marker. A proportion of transformants possessed plasmids that had integrated onto the yeast genome, whilst a class of transformants contained plasmids that were able to replicate autonomously and had higher transformation efficiency (Struhl, Stinchcomb *et al.* 1979). The sequences present on these plasmids became known as autonomously replicating sequences (ARS) (Stinchcomb, Struhl *et al.* 1979). ARS are made up of multiple elements of DNA called A and B elements which are respectively responsible for origin function and the frequency of origin firing (Aparicio, Weinstein *et al.* 1997) (Bell 1995) (Newlon and Theis 1993). The A and B elements are made up of a common 12 bp consensus sequence called the autonomous consensus sequence (ACS), which are required for the binding of the six-subunit origin-recognition complex (ORC) (Aparicio, Weinstein *et al.* 1997) (Bell 1995).



**Figure 1.7 Origins of replication in yeast**

Structure of replication origins in *S. cerevisiae* and *S. pombe*. The ORC binding sites are shown in red and regions required for origin activity are indicated in brown. Regions labelled I, II and III are AT-rich regions, and a bubble with bidirectional arrows indicates a replication origin. Taken from (Gilbert 2001).

In *Schizosaccharomyces pombe* origins of replication are five to ten times larger (0.5-1 kb) than those in *S. cerevisiae* and furthermore the ARS elements do not possess an ACS. Instead ORC binding relies on the presence of characteristic AT rich islands, and poly-dA-poly-dT tracks can replace large sections of ARS elements (Mechali 2010) (Heichinger, Penkett *et al.* 2006) (Okuno, Satoh *et al.* 1999). These AT-rich islands are bound by an AT-hook protein domain of which there are nine present in

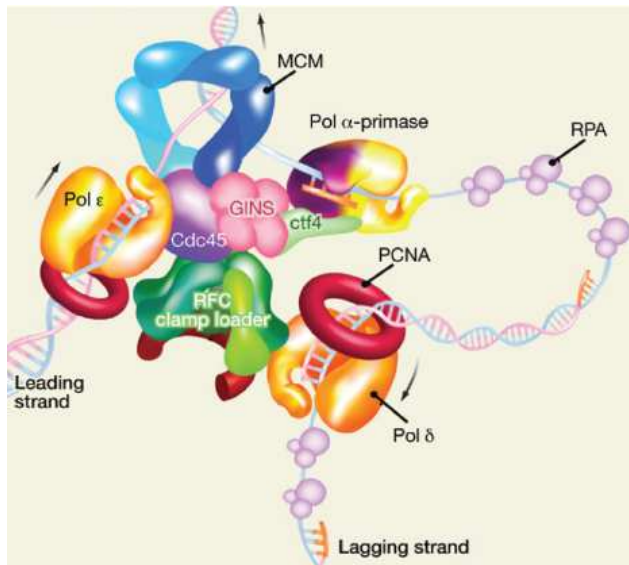


the N-terminus of Orc4 subunit of ORC that is unique to *S. pombe* (Mechali 2010) (Segurado, de Luis *et al.* 2003) (Chuang and Kelly 1999). Orc4 binds specifically to the AT-rich sites without the presence of ATP.

Investigation into human and mouse origins of replication has failed to identify a common consensus sequence, instead a correlation between the presence of both CpG islands and promoter regions with ORC binding has been found (Mechali 2010) (Delgado, Gomez *et al.* 1998). ORC binds to the origin in late S phase early half of G phase and followed by the binding of cell division cycle 6 (Cdc6) and then Cdt1.

During origin licensing the replicative helicase, called the minichromosome maintenance complex (MCM), is recruited (DePamphilis, Blow *et al.* 2006) (Remus, Beuron *et al.* 2009). MCM is made up of a family of six related polypeptides that form a heterohexamer, and is not an ortholog of DnaB despite their functional similarity (Shin and Kelman 2006; Remus and Diffley 2009). MCM loading is facilitated mainly by the MCM loader Cdc6 as well as ORC, MCM9 and Cdt1. MCM is chaperoned to the origin by Cdt1 and duplicate Cdt1:MCM heptamers are loaded onto dsDNA facilitated by ATP hydrolysis. (DePamphilis, Blow *et al.* 2006) (Remus, Beuron *et al.* 2009).

In *S. pombe* the two crucial licensing factors are Cdc18 and Cdt1, the latter of which is essential for origin licensing and is targeted by the Cdc10 transcription factor. Both Cdc18 and Cdt1 are loaded onto the ORC-bound origin by ORC and Sap1, Cdc18 and Cdc1 in turn load the MCM helicase, thus forming the pre-replicative complex and licensing the origin for replication. The transition of the pre-replicative complex to the pre-initiation complex is controlled by the cell cycle through cyclin-dependent kinase (CDK) and Dbf4-dependent protein kinase (DDK) phosphorylation (Sclafani and Holzen 2007).



**Figure 1.8 The eukaryotic replisome**

The structure of the eukaryotic replisome, adapted from (Yao and O'Donnell 2010). See text for further details.

MCM is a highly conserved eukaryotic DNA helicase (Aparicio, Weinstein *et al.* 1997) made up of six closely related MCM2-MCM7 proteins that have each shown to be essential for DNA replication initiation and elongation in *S. cerevisiae* (Moyer, Lewis *et al.* 2006) (Forsburg 2004). MCM forms ring shaped hexamers around DNA at replication origins, each MCM subunit possesses an AAA+ ATPase domain and a conserved MCM box (Remus, Beuron *et al.* 2009). Once loaded onto dsDNA MCM is able to slide passively however in the prereplicative complex MCM remains inactive until S phase where it is activated by CDK and DDK phosphorylation. These two kinases in turn are regulated by the binding of their activating subunit cyclin and Dbf4 respectively, the number of CDKs in metazoans varies greatly (Sclafani and Holzen 2007) (Remus, Beuron *et al.* 2009) (Evrin, Clarke *et al.* 2009).

MCM is regulated by CDK and DDK phosphorylation during late G1 phase on all six subunits except MCM5, leading to the activation of the prereplicative complex and initiation of DNA replication by melting origin DNA. In addition two proteins essential for MCM helicase activity, Cdc45 and the GINS complex (comprising four subunits Sld5, Psf1, Psf2 and Psf3) are recruited to the prereplicative complex, forming the Cdc45/MCM2-7/GINS (CMG) complex (Sclafani and Holzen 2007) (Remus, Beuron *et al.* 2009) (Evrin, Clarke *et al.* 2009) (Pacek, Tutter *et al.* 2006). The binding of Cdc45 and the GINS complex to the MCM helicase forms the CMG complex that has been shown to be the core of helicase activity at the replication fork. It is unclear whether the binding of Cdc45 and/ or GINS to MCM results in a

conformational change, converting MCM to an active form, or if they have a direct role in the unwinding of dsDNA. But what is clear is that alone MCM has a very weak helicase activity which becomes more effective when MCM is bound by Cdc45 and GINS (Moyer, Lewis *et al.* 2006) (Evrin, Clarke *et al.* 2009; Remus, Beuron *et al.* 2009). As a result of MCM helicase activity, topological problems including positive supercoiling and catenation arise which can be resolved by the action of topoisomerase I and II (Diffley and Labib 2002) (Moldovan, Pfander *et al.* 2007).

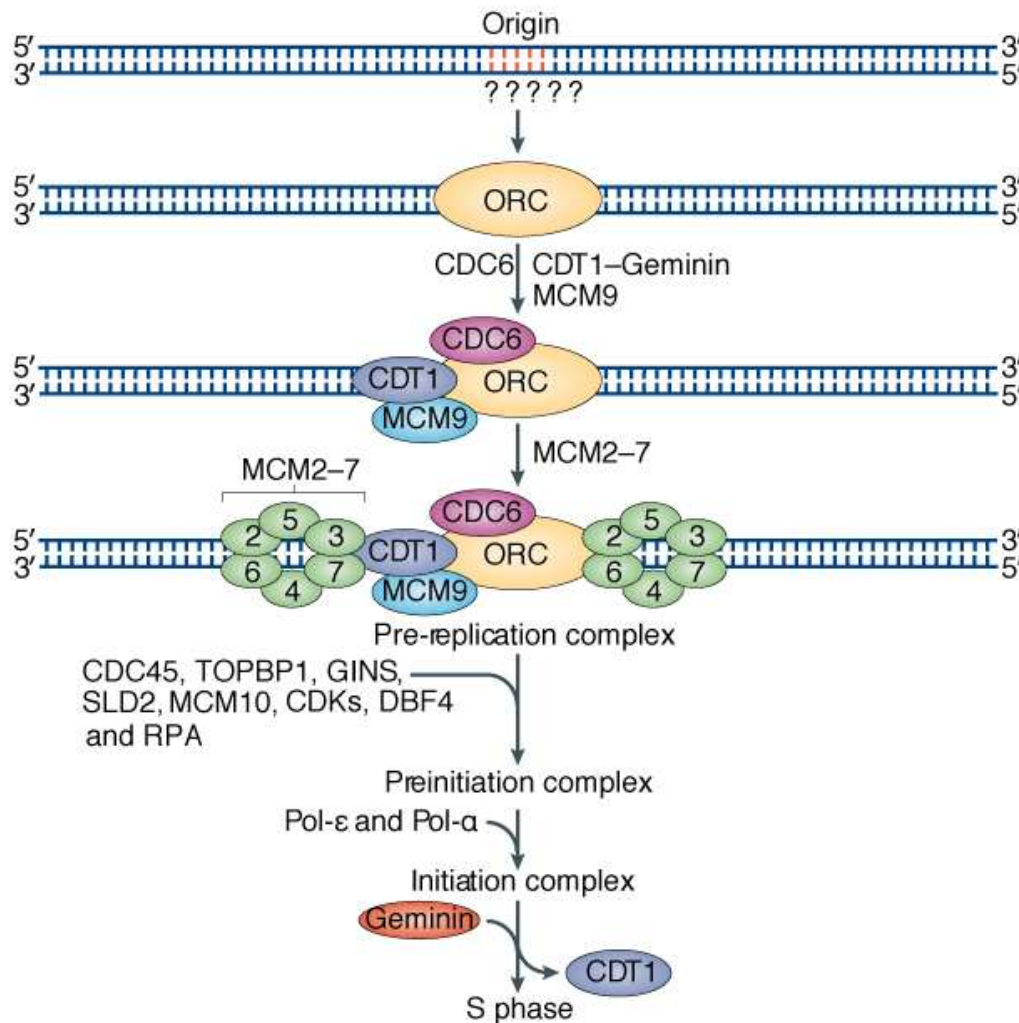
Prior to the recruitment of Cdc45 and GINS, the MCM complex interacts and loads MCM10, which is vital for the loading of Cdc45, GINS, Replication Protein A (RPA) and DNA polymerase  $\alpha$  (Zhu, Ukomadu *et al.* 2007) (Aparicio, Weinstein *et al.* 1997) (Mimura and Takisawa 1998) (Ricke and Bielinsky 2004) (Pacek, Tutter *et al.* 2006) (Gambus, Jones *et al.* 2006) (Takayama, Kamimura *et al.* 2003).

The formation of the CMG complex is an important step in the conversion of the prereplication complex to the preinitiation complex and is closely regulated by CDK and DKK phosphorylation (Sclafani and Holzen 2007). The ssDNA generated by the unwinding activity of the CMG complex is bound by replication protein A (RPA), a ssDNA-binding protein equivalent to SSB in bacteria (Kubota, Takase *et al.* 2003). The binding of RPA to the unwound DNA origin protects it from nuclease degradation and from forming secondary structures that would inhibit the DNA replication process. The structure and mechanism of how RPA performs this role will be discussed in greater detail later.

The progression from the prereplicative complex to the preinitiation complex is tightly controlled by S-phase-promoting CDKs (Mimura and Takisawa 1998). Although CDK regulation is common in metazoans and both budding and fission yeast, differences in how they put it into effect can be seen. For example, instead of the negative CDK regulation of Cdc6 which leads to the ubiquitin-mediated degradation at the end of G1 phase in both budding and fission yeast, in metazoans Cdc6 remains stable upon export from the nucleus at the onset of S phase. Furthermore metazoans have an additional control mechanism through the protein geminin, which binds and inhibits Cdt1 until it is targeted for ubiquitin-mediated proteolysis by the anaphase

promoting complex/cyclosome (APC/C), whose activation is also controlled by CDKs (Diffley and Labib 2002).

Once the pre-initiation complex is formed, the licensing of the origin to replicate is completed. The preinitiation complex binds and loads DNA polymerase  $\alpha$  and DNA polymerase  $\epsilon$  to form the initiation complex through protein:protein interactions with Cdc45 and/or RPA (Diffley and Labib 2002). One method of control is through Cdc45 loading of DNA polymerase  $\alpha$  onto the chromatin (Mimura and Takisawa 1998). However, the DNA polymerases are unable to initiate DNA replication *de novo* and require the DNA primase, comprising of PriL and PriS subunits, to synthesis short RNA primers (6-10 nucleotides long) that the DNA polymerase can then extend. (Diffley and Labib 2002) (Moldovan, Pfander *et al.* 2007) (Frick and Richardson 2001).



**Figure 1.9 The assembly of the pre-replication complex at the origin in eukaryotes**

Taken from (Mechali 2010). The origin is recognized and bound by the origin recognition complex (ORC) that in turn recruits cell division cycle 6 (cdc6), Cdt1 (bound by geminin) and MCM9 which load at least two hexameric minichromosome maintenance complexes (MCM) forming the pre-replication complex. Cdc45 and the GINS complex form the CMG complex with MCM involving MCM10, Ctf4, Dbf4 kinase and cyclin dependent kinases (CDKs). Topoisomerase 2 binding protein 1 (TOPBP1) and replication protein A (RPA) are also recruited, the latter of which assists the loading of Pol  $\alpha$  and  $\epsilon$  with cdc45 to form the initiation complex. On entering the S phase of the cell cycle Cdt1 is released by geminin and consequently degraded.

Bidirectional synthesis of dsDNA leads to the continuous synthesis of the leading strand but discontinuous replication of the lagging strand, resulting in Okazaki fragments at 150-200 bp intervals. To ensure DNA polymerase  $\alpha$  initiates every

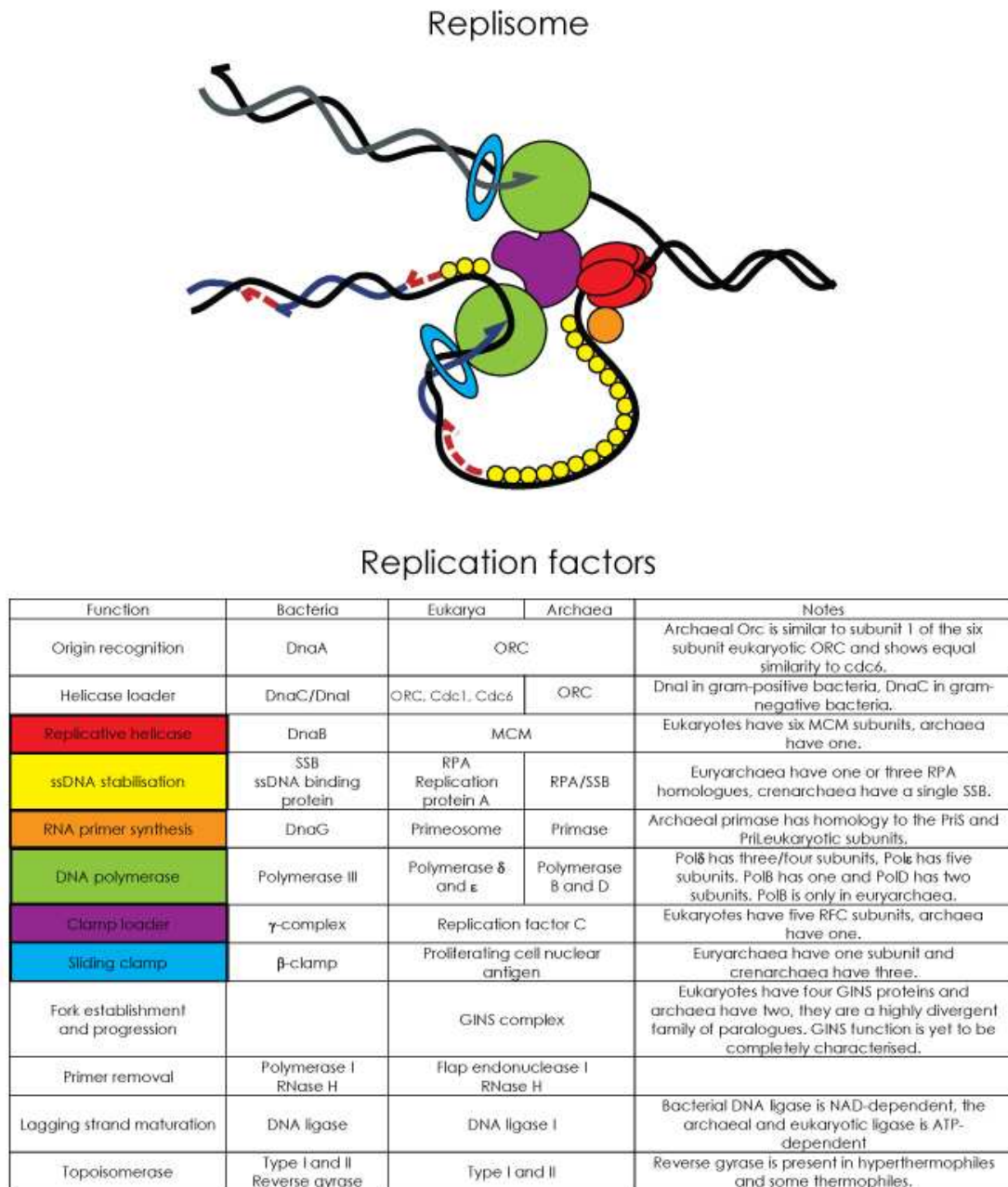
interval, co-ordination of lagging strand synthesis with helicase activity is critical. However, for the continuous synthesis of the leading strand DNA polymerase  $\alpha$  is inappropriate due to its lack of processivity, instead DNA polymerase  $\delta$  and  $\epsilon$  are more suitable, though their specific roles are yet to be fully identified. To assist DNA polymerases, the proliferating cell nuclear antigen (PCNA), a processing factor, is loaded onto the DNA by the 'clamp loader' replication factor C (RFC). PCNA is a member of the sliding  $\beta$  clamp family and although there is no sequence similarity between bacterial and eukaryotic family members, their structure is almost superimposable. RFC recognizes the primer:template junction and catalyses the loading of PCNA here. PCNA plays an important role in DNA replication by tethering replicative enzymes to the replication fork, including DNA polymerases  $\alpha$ ,  $\delta$  and  $\epsilon$ , FLAP endonuclease FEN1 and DNA Lig1 (Moldovan, Pfander *et al.* 2007). Eukaryotic RFC is made up of five homologous subunits (Indiani and O'Donnell 2006) (Cullmann, Fien *et al.* 1995). Once loaded the PCNA encircles the DNA and is able to slide passively whilst stabilizing the DNA polymerases binding and nascent strand synthesis. (Diffley and Labib 2002) (Moldovan, Pfander *et al.* 2007). The presence of PCNA on the DNA leads to the recruitment of the more processive replicative DNA polymerase  $\delta$  or  $\epsilon$ , which take over from DNA polymerase  $\alpha$ . Both DNA polymerase  $\delta$  and  $\epsilon$  possess a 3'-5' exonuclease activity that significantly reduces the stable misincorporation of nucleotides. The mechanism by which the Okazaki fragments are processed to form one nascent lagging strand has been under long debate. When the replicative helicase reaches the end of a previous Okazaki fragment, synthesis persists resulting in a partially displaced Okazaki fragment flap structure. To resolve this and remove the RNA primers, it was originally thought the sequential action of RNase H2 and the FLAP endonuclease FEN1 were responsible, then a second mechanism involving the nuclease/helicase Dna2 and FEN1. The most recent model involves the DNA polymerase  $\delta$  and FEN1 action in a processive complex with PCNA, where the 3'-5' exonuclease activity of DNA polymerase  $\delta$  or  $\epsilon$  and FEN1 cooperate in producing a ligatable nick which DNA ligase I is able to seal. (Diffley and Labib 2002; Moldovan, Pfander *et al.* 2007) (Garg, Stith *et al.* 2004).

DNA replication termination arises when two opposing replication forks meet and the two nascent DNA strands are ligated together, and generally occurs in a random

manner in between replication origins (Diffley and Labib 2002) (Codlin and Dalgaard 2003). However site-specific fork barriers resulting in replication termination can be seen in specific genetic loci, for example in *S. cerevisiae* the 3' terminus of rDNA possesses a polarity that ensures unidirectional replication of the rDNA (Brewer and Fangman 1988). Replication fork barriers prevent the collision of replication and transcription machinery in highly transcribed regions including ribosomal operons (Biswas and Bastia 2008).

### 1.3.3 Archaeal DNA replication

Despite similarity between the circular archaeal chromosomal structure and that of bacteria, most archaea possess multiple DNA replication origins coupled with a similar enzymatic process of DNA replication to that of eukaryotes.

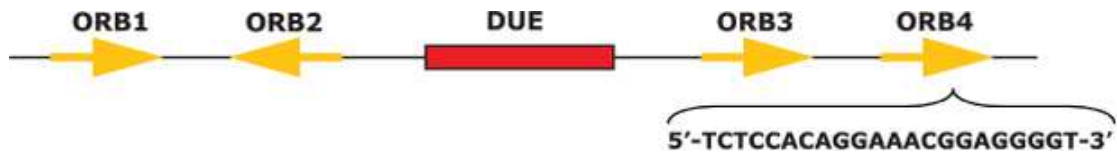


**Figure 1.10 The replisome and replication factors from the three domains of life**

The replisome is depicted above with a table of replication factors from bacteria, eukaryotes (eukarya) and archaea. For colours refer to the replisome figure. Taken from (Hawkins 2009).



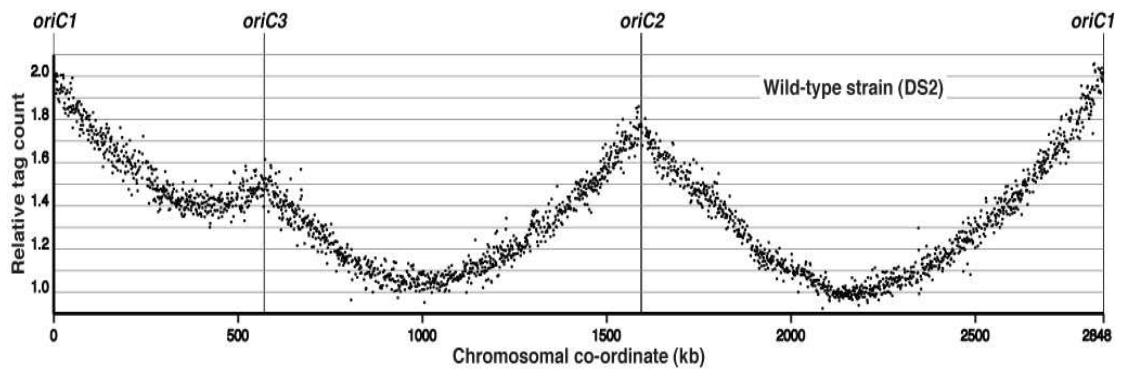
Archaeal genome sequence analysis has revealed that origins of replication are made up of an AT-rich region called a duplex unwinding element (DUE), flanked by extended repeat sequences named the origin recognition box (ORB) elements (Gaudier, Schuwirth *et al.* 2007).



**Figure 1.11 Archaeal origin structure**

The structural organization of the *Aeropyrum pernix* origin of replication, adapted from (Gaudier, Schuwirth *et al.* 2007). The ORB elements, depicted by yellow arrows, flank the DUE element. The top strand sequence of ORB4 is shown.

The number of origins in archaea varies, single origins are seen in *Pyrococcus abyssi* (Myllykallio, Lopez *et al.* 2000) and *Archaeoglobus fulgidus* (Maisnier-Patin, Malandrin *et al.* 2002). Multiple origins are seen in *Sulfolobus acidocaldarius* (Duggin, McCallum *et al.* 2008), *Sulfolobus solfataricus* (Lundgren, Andersson *et al.* 2004), *Halobacterium* (Coker, DasSarma *et al.* 2009) and *Haloferax volcanii*, which was originally thought to have two replication origins on the main chromosome (Norais, Hawkins *et al.* 2007). Recent DNA replication marker frequency analysis (MFA) by new generation sequencing of early exponential *H. volcanii* cells relative to stationary phase cells has revealed a third active origin of replication, as shown in Figure 1.12. Interestingly origin independent replication is observed in *H. volcanii* deleted for all three replication origins at a comparable rate to that of wild type, with MFA confirming the absence of a dormant origin (Hawkins, Retkute *et al.* Submitted).



**Figure 1.12** Chromosome replication profile of *Haloferax volcanii*

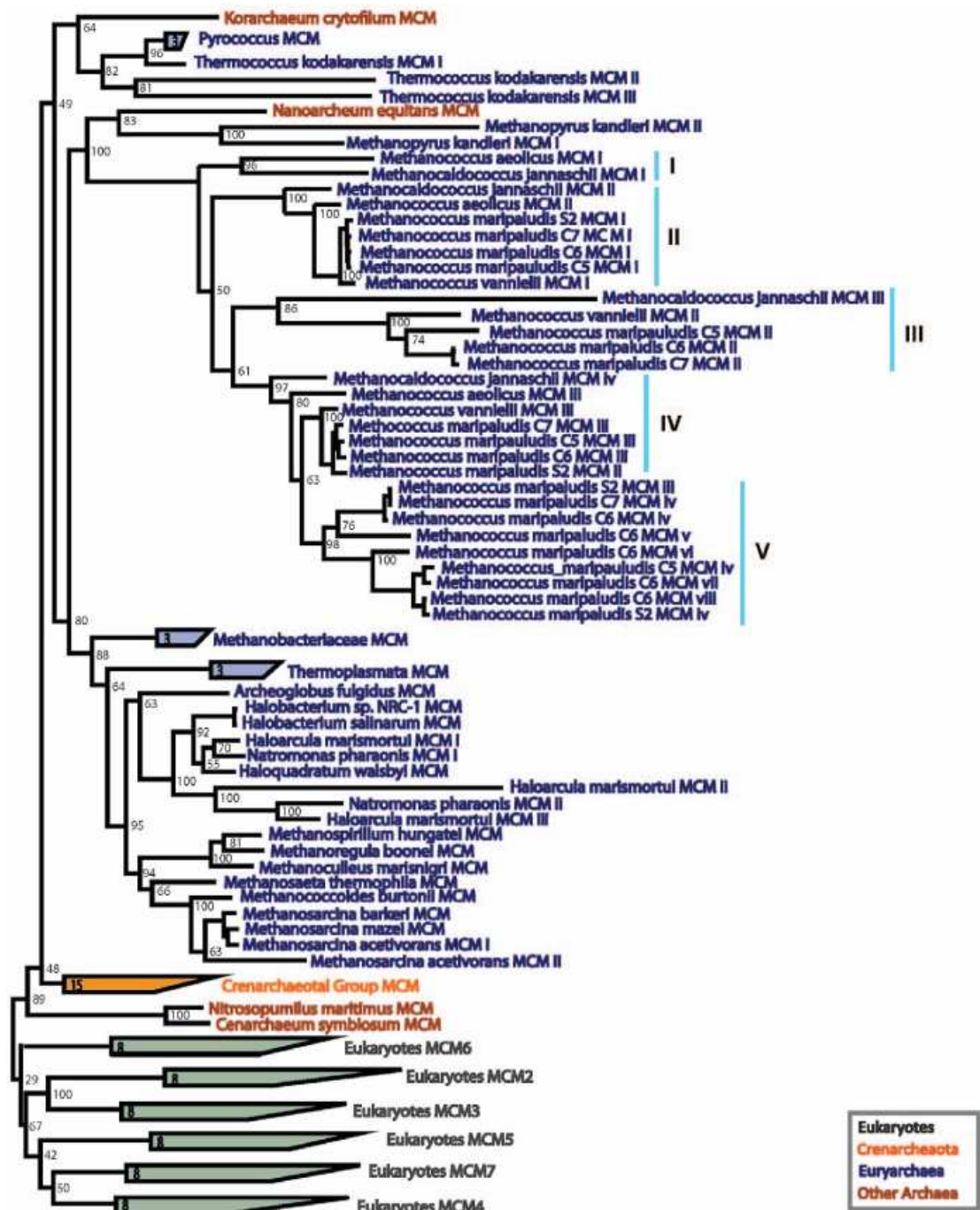
The tag counts of the main chromosome of *Haloferax volcanii* are analysed by MFA. Sharp peaks indicate discrete origins and smooth valleys denote random termination of DNA replication. Adapted from (Hawkins, Retkute *et al.* Submitted).

Archaeal origins of replication are recognized and bound by the archaeal origin recognition protein Orc1/Cdc6, which occurs in multiple homologs. *Sulfolobus solfataricus* possesses three Orc1/Cdc6 homologs (Barry and Bell 2006) and *H. volcanii* possesses sixteen Orc1/Cdc6 homologs, whose genes are distributed amongst different replicons (Norais, Hawkins *et al.* 2007). The origin recognition protein Orc1/Cdc6 is denominated as such as it shows homology to both the eukaryotic origin recognition protein Orc1 and the helicase loader Cdc6, yet it is still to be determined if it functions as a origin recognition complex or a helicase loader, or both (Barry and Bell 2006; Akita, Adachi *et al.* 2010). Orc1/Cdc6, a member of the AAA+ ATPase family, does not use ATP hydrolysis to bind the ORB elements, instead this activity is believed to load the MCM helicase in the formation of the prereplicative complex (Barry and Bell 2006) (Kasiviswanathan, Shin *et al.* 2005).

The role of loading the archaeal replicative helicase MCM, a homolog of the eukaryotic MCM, is believed to lie with Orc1/Cdc6, due to the absence of an archaeal homolog to either the eukaryotic or bacterial helicase loader Cdt1 or DnaC respectively. Phylogenetic analysis of archaeal MCM homohehexamer subunits, shown in Figure 1.13, shows great variance in the number of distinct subunits present in different archaeal species, arising from gene duplication events. For example *Methanococcus maripaludis* C5 possesses eight different MCM subunits. This contrasts highly with eukaryotes, where all species possess the same six MCM subunits that form a heterohexamer. This suggests that there is no correspondence

between the archaeal MCM subunits and the eukaryotic subunits in functioning as heterohexamer. It has been suggested that the archaeal MCM proteins have become structurally refined and specialized in protein:protein interactions (Chia, Cann *et al.* 2010). Recent studies into *Methanococcales* MCM subunits has revealed an ancient duplication, pre-dating the divergence of *Methanococcales* into individual species, resulting in two major groups of MCM subunits (A and D) (Walters and Chong 2010).

This has been observed in the archaeon *Thermococcus kodakarensis*, which possesses three MCM homologs, MCM1, MCM2 and MCM3. Although all three MCMs have been shown to be capable of binding ssDNA, only MCM2 and 3 display typical archaeal MCM activity as well as forming hexameric structures. MCM3 is believed to be the predominant replicative helicase and essential due to the inability to generate a  $\Delta mcm3$  deletion mutant. Both  $\Delta mcm1$  and  $\Delta mcm2$  deletion mutants can be generated, and in addition to co-purification of MCM1 and 2 with DNA polymerase B and D, PCNA1 and MutS homologs (with MCM2 only), suggests MCM1 and 2 may have specialized functions in DNA repair or recombination (Pan, Santangelo *et al.* 2011).

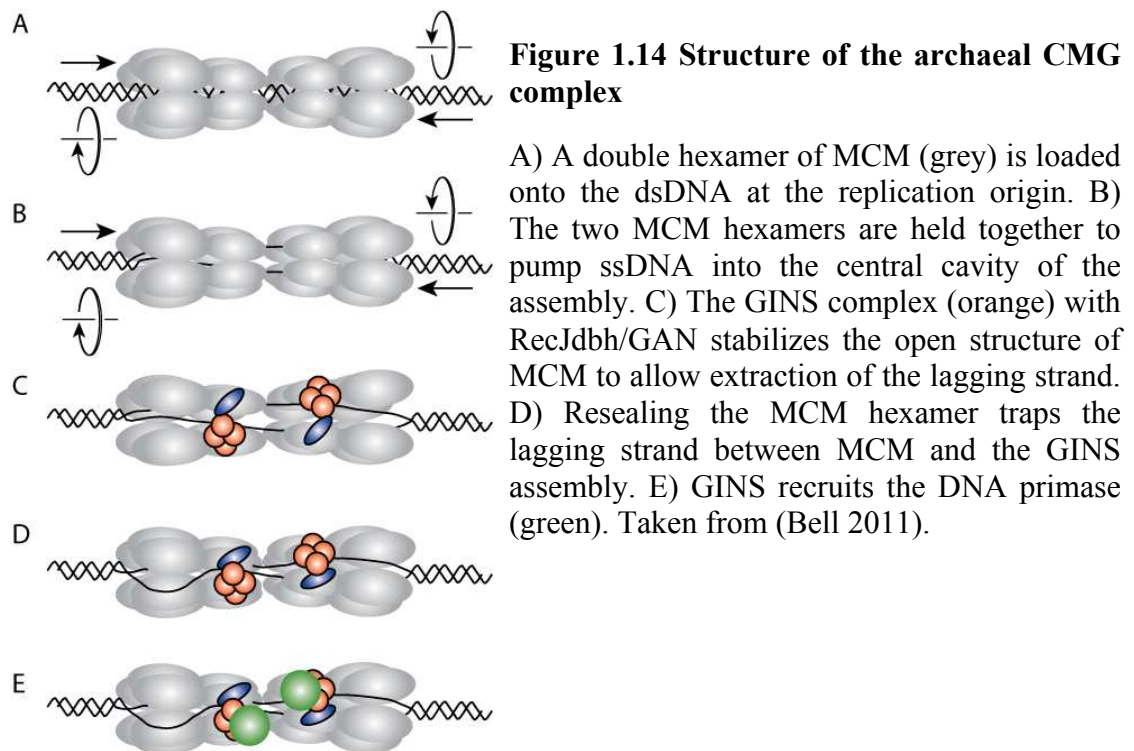


**Figure 1.13 Phylogenetic analysis of archaeal and eukaryotic MCM subunits**

Taken from (Chia, Cann *et al.* 2010). MCM subunits arising from gene duplication events labelled I-V according to the phylogeny.

The presence of GINS is conserved from eukaryotes to archaea, however the structure differs to the eukaryotic tetramer (which is made up of Sld5, Psf1, Psf2 and Psf3 subunits). Some archaeal GINS do exist as heterotetrameric complexes made up of two subunits of Gins15, (most closely related to eukaryotic Psf1 and Sld5), and

Gins23, (most closely related to eukaryotic Psf2 and Psf3), as seen in *Thermococcus kodakaraensis*. However, many euryarchaea only possess one homolog, for example *Thermoplasma acidophilum* has only the Gins15 homolog and *Haloferax volcanii* possesses a single gene *ginS* that is essential. The archaeal GINS forms the CMG complex with MCM, as seen in eukaryotes, but with GINS-associated nuclease (GAN) also known as RecJdbh, a RecJ family nuclease with a homologous ssDNA binding domain to that of RecJ. These were identified in *T. kodakaraensis* and *Sulfolobus solfataricus* respectively (Bell 2011). As seen in eukaryotes, it is proposed that MCM loads onto dsDNA at replication origins as a head-to head double hexamer, shown in Figure 1.14. The presence of RecJdbh/GAN and GINS would facilitate the temporary opening of the MCM hexamer to allow extrusion of the lagging strand. As the duplex DNA is unwound, the leading strand has been shown to pass through the centre of the archaeal MCM whilst the lagging strand is delivered, by the CMG complex, to the DNA primase. This theory is supported by observed protein:protein interaction between Gins23 and the DNA primase (Macneill 2009) (Bell 2011).



The ssDNA resulting from MCM helicase activity is bound a single-stranded DNA binding protein as in bacteria and eukaryotes. Overall, archaea show a greater similarity to the eukaryotic RPA than the bacterial SSB. However, the number of RPA

proteins present and the complexes they form differs widely from species to species and will be discussed in greater detail later (Barry and Bell 2006). Once the ssDNA is bound by RPA, the DNA primase is then able to associate and synthesise short RNA primers required for DNA synthesis initiation (Grabowski and Kelman 2003). The archaeal primase, a homologue of the eukaryotic DNA primase, synthesises RNA primers 7 -10 bp in length (Barry and Bell 2006) (Lao-Sirieix, Nookala *et al.* 2005).

The RNA primer is extended by the DNA polymerases and thus DNA synthesis is initiated. Archaea possess two families of DNA polymerases; family B and family D, the latter of which is only found in euryarchaea and thaumarchaeota (Barry and Bell 2006) (Spang, Hatzenpichler *et al.* 2010). All eukaryotic DNA polymerases ( $\alpha$ ,  $\delta$ ,  $\epsilon$ ) belong to the family B polymerases that are found in all archaea and possess a strong 3'-5' proofreading activity (Grabowski and Kelman 2003). Characteristic activity of archaeal family B polymerases is their ability to recognize and bind tightly to uracil and hypoxanthine, stalling replication and preventing a CG to a TA base pair conversion (Killelea, Ghosh *et al.* 2010).

Family D polymerases, only found in euryarchaea and thaumarchaeota, have been shown to be essential and are made up of two subunits DP1 and DP2, the former of which possesses 3'-5' exonuclease activity and is the proof-reading subunit (Spang, Hatzenpichler *et al.* 2010) (Barry and Bell 2006). A study of *P. furiosus* B and D polymerases supports the model where PolB is responsible for DNA synthesis on the leading strand and PolD the lagging strand (Barry and Bell 2006) (Henneke, Raffin *et al.* 2000).

To enable the DNA polymerase on the leading strand to continuously synthesize without falling off the template processive accessory factors including PCNA and RFC are required. PCNA has been described as a molecular 'tool belt', encircling DNA and tethering interacting proteins to the DNA substrate. The archaeal PCNA is a homotrimer, unlike the heterotrimeric eukaryote homolog, and exists as a single homolog in all euryarchaea. With the exceptions of *Thermococcus kodakarensis* and crenarchaea where multiple PCNA homologs exist, (for example both *Aeropyrum pernix* and *S. solfataricus* have three homologs), once again displaying the diversity of

archaea. Investigation into *Sulfolobus* PCNA revealed that PCNA interacting proteins including FEN1, DNA polymerase B1 and DNA ligase I interact with specific PCNA subunits via the PCNA-interacting protein (PIP) motif (Indiani and O'Donnell 2006) (Castrec, Rouillon *et al.* 2009). EM reconstruction analysis of the *P. furiosus* RFC-PCNA-DNA ternary complex has shown that the PCNA undergoes a conformational change during the process of clamp-loading (Miyata, Suzuki *et al.* 2005) (Kazmirski, Zhao *et al.* 2005). As in eukaryotes, PCNA is loaded onto DNA by RFC, the clamp loader (Barry and Bell 2006) (Dionne, Nookala *et al.* 2003).

Archaeal RFC, a pentamer consisting of a large subunit (RFCL) and four identical small subunits (RFCS), is responsible for catalysing the opening and assembling of PCNA onto dsDNA- ssDNA junctions such as a primer- template junction. Both RFCS and RFCL possess seven of the eight highly conserved RFC motifs known as box II-VIII in both human and yeast (Cullmann, Fien *et al.* 1995). Archaeal RFC shares an average of 40% sequence identity with eukaryotic RFC, with the RFCL subunit being the more divergent (Cann and Ishino 1999). In each haloarchaea genome, there are three genes encoding distinct RFC proteins that have been shown to be essential in *H. volcanii* (Macneill 2009). The binding of ATP by the *A. fulgidus* RFC has been shown to induce an open conformational change of the RFC that is believed to correspond to the opening of the PCNA structure by RFC during clamp-loading (Seybert, Singleton *et al.* 2006), after which ATP hydrolysis allows the release of PCNA onto the DNA. Once loaded onto the DNA, the PCNA is able to slide passively, assisting the DNA polymerase in processive DNA synthesis during the elongation phase of DNA replication (Indiani and O'Donnell 2006). DNA synthesis on the lagging strand results in a series of Okazaki fragments separated by RNA primers. The RNA primers are thought to be removed by the flap endonuclease FEN1, which possesses a 5'-flap endonuclease and 5'-3' endonuclease activities, and RNase H2, which degrades the RNA primer. The resulting gap is sealed by the DNA ligase LigA, which is localized to the nick by PCNA, resulting in the Okazaki fragments on the lagging strand to form one nascent strand. (Barry and Bell 2006) (Grabowski and Kelman 2003) (Macneill 2009). *H. volcanii* possesses a NAD<sup>+</sup>-dependent DNA ligase LigN, in addition to a ATP-dependent ligase, LigA. Although neither LigN nor LigA are singularly essential, a genomic deletion of both is

synthetically lethal, suggesting the DNA ligases have an overlapping function (Zhao, Gray *et al.* 2006).

Very little is known of termination of DNA replication, however in *Sulfolobus solfataricus* fusion of two DNA replication forks results in termination of DNA replication, as opposed to site-specific termination. Whether collision of replication forks is random or occurs at specific sites is unknown (Duggin, Dubarry *et al.* 2011). In *H. volcanii* MFA has also revealed the use of broad termination zones, shown by smooth valleys in Figure 1.12 (Hawkins, Retkute *et al.* Submitted).

## 1.4 DNA repair

DNA repair occurs throughout the cell cycle and generates ssDNA in the process, which is bound by the ssDNA-binding protein RPA. The binding of RPA prevents the ssDNA from forming secondary structures such as hairpins, that would prevent DNA repair proteins from accessing the DNA, in addition to protecting the ssDNA from being degraded by nucleases that are required in DNA repair processes, the ssDNA is bound by RPA. RPA interacts with DNA repair proteins whilst bound to ssDNA, and once the presence of RPA bound to ssDNA is no longer desirable, RPA is displaced via protein:protein interactions. For example, in eukaryotes RPA is displaced by the recombination mediator Rad52, to allow the loading of the recombinase Rad51 onto the ssDNA during homologous recombination. Therefore to comprehend the role of RPA in the cell, the DNA repair processes must first be understood.

During the cell cycle DNA is susceptible to damage during metabolism or by exposure to mutagens or carcinogens, which can result in a variety of DNA lesions. These include double or single-stranded DNA breaks, the hydrolysis of purine bases resulting in abasic sites; deamination, the oxidation of amine groups on the bases to aldehydes; isomerisation of bases; DNA adducts including cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts, as well as inter and intra-strand DNA cross-links. All of which pose a threat to the genome integrity and the ability of the cell to replicate faithfully.



#### **1.4.1 Nucleotide excision repair**

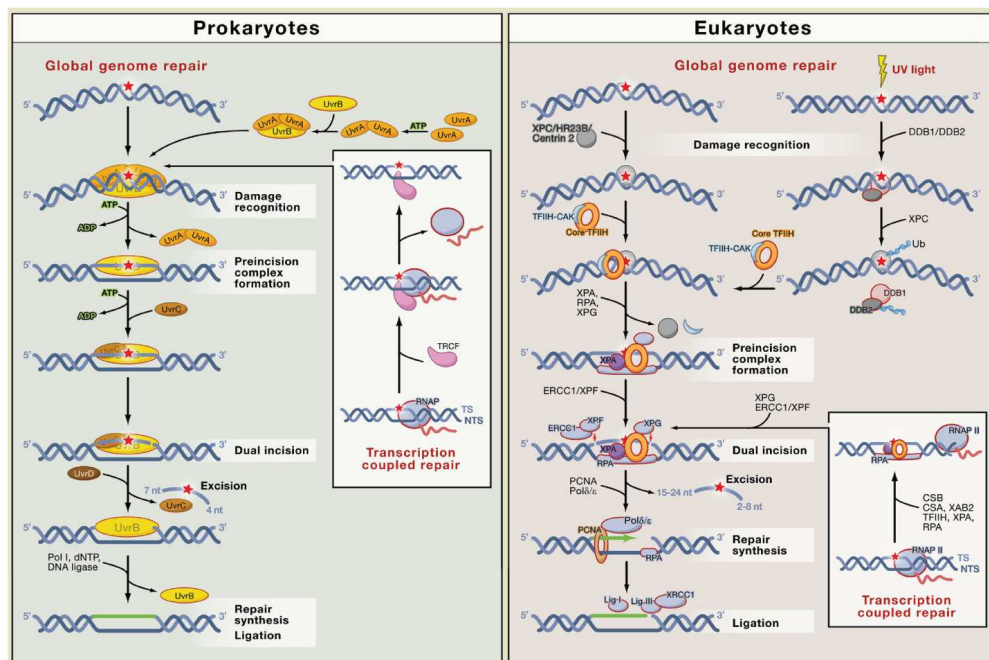
The nucleotide excision repair (NER) pathway is used to remove bulky lesions that distort the DNA helix, including UV-induced CPDs and pyrimidine (6-4) pyrimidone photoproducts (6-4PPs). The NER system is essentially a “cut-and-patch” one and involves detection of the lesion, DNA unwinding and incision, removing a short stretch of nucleotides containing the lesion after which the resulting gap is filled using the non-damaged DNA strand as a template (Van Houten, Croteau *et al.* 2005) (Rouillon and White 2011). The inability for a cell to perform NER correctly has been shown to result in rare autosomal recessive human disorders such as Xeroderma Pigmentosum, Trichothiodystrophy and Cockayne syndrome, all of which show sensitivity to sunlight (Fousteri and Mullenders 2008). The NER pathway can be divided into two mechanisms: global genome repair and transcription-coupled repair, which differ mainly in the molecular mechanism employed to recognize the DNA lesion, and that transcription coupled repair involves DNA repair during transcription, on the transcribed strand (Guo, Tang *et al.* 2010).

#### **1.4.2 Global Genome repair**

As suggested by its name global genome repair (GG-NER) recognises and repairs DNA lesions throughout the genome on both DNA strands. The bacterial GG-NER mechanism involves the UvrABC system, where the UvrAB complex scans DNA for adducts until UvrA recognizes a lesion. UvrA transfers the DNA lesion to the DNA binding domain of UvrB, where UvrB verifies the lesion using its  $\beta$ -hairpin structure to separate the two DNA strands. UvrA then dissociates via ATP hydrolysis allowing UvrC to make a 3' and a 5' incision resulting in a 12 nucleotide fragment. The DNA helicase II (UvrD) releases UvrC and the 12 nt fragment. The resulting gap is filled by DNA polymerase I and is joined to the parent DNA by DNA ligase I (Van Houten, Croteau *et al.* 2005).

In eukaryotes the heterodimeric complex UV-DDB, XPE (DDB1-DDB2-containing E3-ubiquitin ligase complex) and XPC-RAD23 recognise and bind the DNA lesion due to the helical distortion. TFIIH, XPA and RPA are recruited to open and unwind the helix, allowing NER lesion verification. Here the helicase action of the XPD and the ATPase activity of the XPB subunits unwinds the dsDNA, and RPA plays an important role in stabilizing the resulting ssDNA, to allow DNA incisions to be made.

The XPF-ERCC1 and XPG endonucleases make 3' and 5' incisions, respectively, either side of the NER lesion. The resulting 25-30 nucleotide gap is then filled by the action of DNA polymerase  $\delta$  and  $\epsilon$ , which as in DNA replication are tethered to DNA by the proliferating cell nuclear antigen (PCNA), which in turn is loaded by the clamp loader RFC. DNA synthesis fills in the gap using the non-damaged DNA strand as a template. XRCC1/DNA ligase III seals the nick in addition to DNA ligase I, which participates in NER during the S phase of the cell cycle, joining the newly synthesised DNA to the parent DNA (Fousteri and Mullenders 2008) (Palomera-Sanchez and Zurita 2011).



**Figure 1.15 Nucleotide excision repair (NER)**

Nucleotide excision repair pathway in bacteria (left hand box), UvrA,B and C are responsible for the recognition and excision of the damaged base and require both ATP binding and hydrolysis. After excision, by UvrC, UvrD (DNA helicase II) facilitates the release of the excised damaged oligonucleotide as well as the dissociation of UvrC. DNA polymerase I fill in the resulting gap and the remaining nick is sealed by DNA ligase. Nucleotide excision repair in eukaryotes (right hand box) where the complementation group of xeroderma pigmentosum (XP), designated XPA to XPG, function. XPC/HR23B/Centrin2 and DDB1/DDB2 (UV-DDB) function in the recognition of the damaged base. Ubiquitination of NER proteins provides control through repression or stimulation, for example DDB2 and XPC respectively. The transcription factor IIH (TFIIH), XPA, RPA and XPG drive the formation of the bubble allowing XPG and ERCC1/XPF excision of the lesion from the genome. DNA polymerase  $\delta$  or  $\epsilon$ , PCNA, RPA and DNA ligase I or XRCC1 and DNA ligase III are responsible for filling the remaining gap. Insets in both cases show the pathway for transcription-coupled repair, where in bacteria the transcription repair coupling factor (TRCF) is responsible for recruiting UvrA and B and in eukaryotes CSB, an ATP-dependent chromatin-remodelling factor, acts with CSA to facilitate NER proteins access to the damaged base. Taken from (Guo, Tang *et al.* 2010).

### **1.4.3 Transcription-coupled nucleotide excision repair**

Transcription elongation complexes that are stalled due to DNA lesions are removed by transcription-coupled nucleotide excision repair (TC-NER). Fast and effective repair of a DNA lesion that would otherwise inhibit the RNA polymerase II complex is crucial. This pathway specialises in removing CPDs, 6-4PPS and DNA-protein complexes induced by topoisomerase I inhibitors. In *E. coli* the transcription-repair

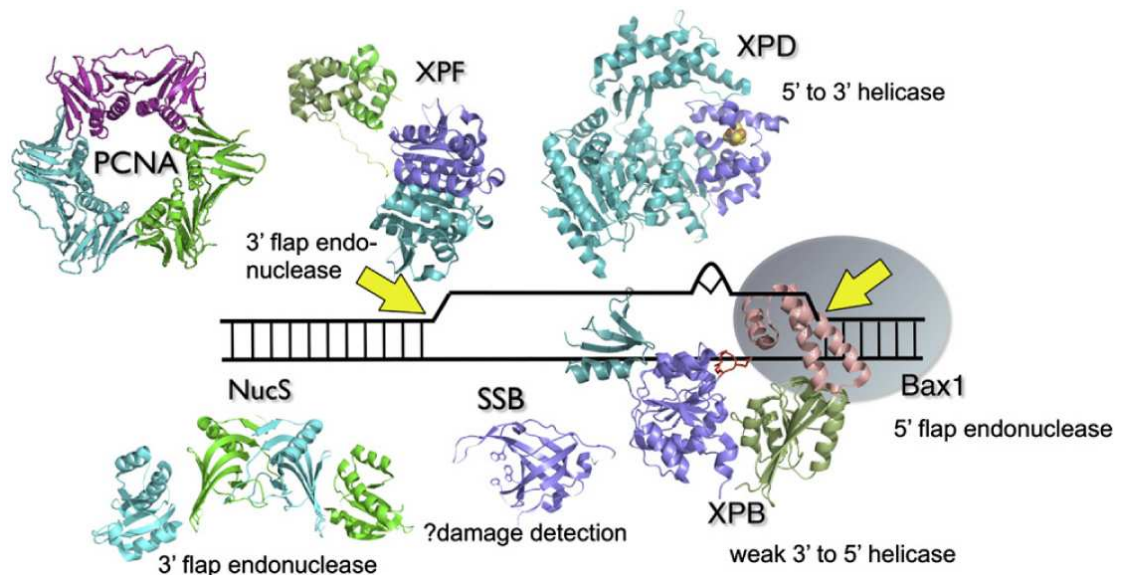
coupling factor (TRCF), encoded by the *mfd* (mutation frequency decline) gene, facilitates the ATP-dependent release of the RNA polymerase and the newly synthesised transcript from the damaged DNA, as well as recruiting members of the NER machinery, in particular UvrA. In budding yeast *S. cerevisiae* the *mfd* counterpart Rad26 is responsible for the recognition of the damaged DNA during transcription.

Defects in TC-NER after exposure to UV-light and chemicals that result in bulky DNA adducts are found in Cockayne's syndrome. Humans possess CSA and CSB, both TRCF and Rad26 counterparts. CSB, a DNA dependent ATPase DNA binding and nucleosome remodelling protein, is able to interact stably with the RNA polymerase II complex in the presence of DNA damage. CSA, an inactive ubiquitin ligase when involved in NER, is recruited by CSB after which the NER machinery is recruited. However unlike bacterial TC-NER, the RNA polymerase is not displaced and is only degraded as a last resort when the DNA lesion cannot be removed and repaired by TC-NER. Here Def1 promotes the recruitment of ubiquitylation machinery that results in the degradation of the RNA polymerases, allowing alternative DNA repair such as (Fousteri and Mullenders 2008).

#### **1.4.4 Archaeal nucleotide excision repair**

Archaea possess homologs to eukaryotic NER proteins, however their distribution is not consistent throughout this domain. The helicases XPB and XPD are both individually absent in *Methanocaldococcus jannaschii* and *Nanoarchaeum equitans* respectively, and both absent in *Methanothermobacter thermautotrophicus* where the bacterial UvrABC system is present instead. Homologs of the bacterial UvrABC proteins are found in haloarchaea as well as homologs of *S. cerevisiae* Rad2, Rad3 and Rad25 (Crowley, Boubriak *et al.* 2006) (McCready and Marcello 2003). *Methanopyrus kandleri* does not possess XPB, XPD or the bacterial UvrABC proteins. In addition a gene encoding the Bax1 5' flap endonuclease, which fulfils the same role as XPG in eukaryotes, is co-transcribed with the *xpb* gene and has been shown to form a heterodimer with XPB (Rouillon and White 2011). The ssDNA-binding protein RPA (depicted as SSB in Figure 1.16) can recognize DNA lesions including mismatches, photoproducts and bulky lesions in *Sulfolobus solfataricus*. Due to the

transient ssDNA produced by the DNA lesions, RPA was able to bind and destabilize them. This has led to the theory that RPA recognizes NER lesions and initiates archaeal GG-NER (Cubeddu and White 2005). The figure below depicts a simplified eukaryotic NER system in archaea, there are many aspects still to be resolved (Rouillon and White 2011).



**Figure 1.16 Archaeal nucleotide excision repair**

A suggested model for crenarchaeal nucleotide excision repair taken from (Rouillon and White 2011). Structure for PCNA taken from *Sulfolobus solfataricus*, XPF from *Aeropyrum pernix*, XPD from *Thermoplasma acidophilum*, NucS from *Pyrococcus abyssi*, SSB (RPA) from *S. solfataricus* and XPB from *Archaeoglobus fulgidus*. The structure of Bax1 is yet to be determined. PCNA has been shown to have a catalytic role during NER, both interacting directly with and stimulating XPF and NucS. The dsDNA region containing the lesion is unwound by XPD and XPB, the former of which is able to bypass the lesion. XPB also interacts with Bax1 to facilitate the unwinding and nicking of the DNA lesion. SSB may have a role in DNA damage detection and the exact role of NucS is yet to be determined.

#### **1.4.5 Base excision repair**

Base excision repair is a multi-step recognition process that allows the distinction between a damaged base arising from oxidative damage, alkylation and deamination, and a non-damaged base, followed by the removal and replacement of the damaged base. The BER process in eukaryotes is essentially equivalent to that of bacteria (Lindahl 2001). The core BER machinery includes a lesion specific DNA glycosylase,

an AP endonuclease or an AP DNA lyase, a DNA polymerase and a DNA ligase (David 2005) (Robertson, Klungland *et al.* 2009).

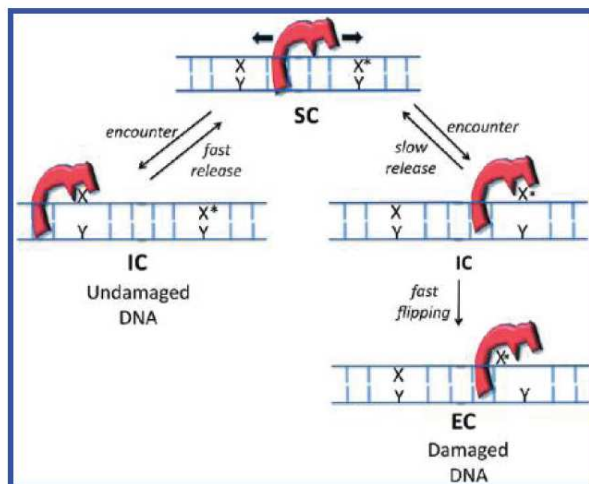
### **DNA Glycosylases**

The role of a DNA glycosylase in BER is to specifically recognise a damaged base as a substrate and catalyse the hydrolysis of the N-glycosidic bond that releases the base and creates an abasic site. DNA glycosylases have specific substrates; the uracil-DNA glycosylase, named UNG in *E. coli* and UDG in humans, is specific for uracil in DNA, arising from deamination of cytosine. The ability of uracil to base-pair with adenine would generate a C-to-T transition mutation if it were not removed by UNG or UDG. The enzyme 8-oxoguanine glycosylase OOG1 in humans and *S. cerevisiae* specifically removes 8-oxoguanine (oxoG). In *E. coli* the FaPy DNA glycosylase MutM (FPG) is responsible for the removal of oxoG and formamidopyrimidine (FaPy). *E. coli* also possesses an adenine DNA glycosylase MutY, whose substrate is a mispaired adenine with guanine or oxoG, which can arise from reactive oxygen species or ionising radiation damage. (Robertson, Klungland *et al.* 2009).

The process in which a DNA glycosylase removes a damaged base is a three-step process as shown in the Figure 1.17. The damaged base is identified by DNA glycosylase in the search complex (SC) which then makes the transition to interrogation complex (IC) to verify the damaged base as a correct substrate. Finally the transition of the DNA glycosylase from IC to excision complex (EC) involves the flipping of the damaged base, positioning the base in the active site of the DNA glycosylase ready for excision (Friedman and Stivers 2010).

## AP Endonucleases

AP endonucleases recognize the AP site and incise 5' of the lesion resulting in a single-stranded break (SSB) with a 5'-deoxyribosephosphate (5'-dRP) end. In *S. cerevisiae* the AP endonucleases are Apn1 and Apn2, of which the former accounts for the majority of AP endonuclease activity (Boiteux and Guillet 2004). In *E. coli* exonuclease III and endonuclease IV are the two major AP endonucleases and in humans HAP/APE (Cunningham, Saporito *et al.* 1986) (Johnson, Torres-Ramos *et al.* 1998).

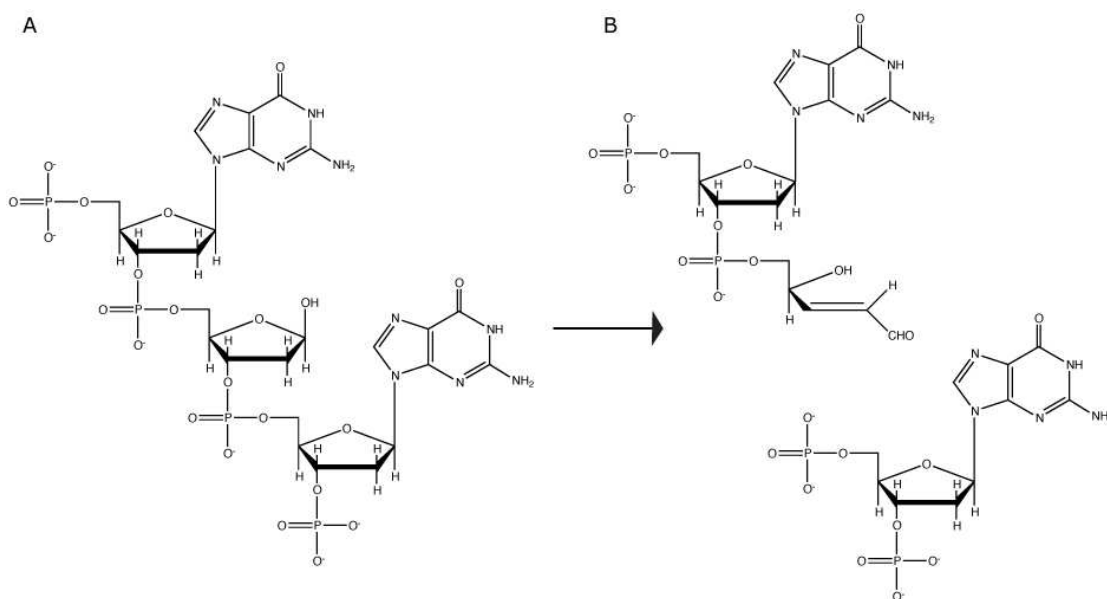


**Figure 1.17 DNA glycosylase complexes during BER**

The DNA glycosylase searches for damaged bases by sliding along DNA in the loose search complex (SC). Transition to the interrogation complex (IC) allows close examination of potential substrate base which if proved to be so triggers final transition to excision complex (EC), involving base flipping of the damaged base. Taken from (Friedman and Stivers 2010)

AP lyases also act on AP sites via a  $\beta$ -elimination reaction, making a 3' incision and yielding a 3' unsaturated aldehydic  $\alpha$ ,  $\beta$ , 4-hydroxy-2-pentanal (3'-dRP) end which can be removed by a 3' phosphodiesterase. AP lyase activity is possessed by many DNA glycosylases including *E. coli* MutM, *S. cerevisiae* Ntg1, Ntg2 and OGG1 and

human NTH1 (Boiteux and Guillet 2004) (Robertson, Klungland *et al.* 2009) (Oyama, Wakasugi *et al.* 2004).



**Figure 1.18  $\beta$ -elimination reaction**

A) AP site upon which the AP lyase makes a 3' incision, resulting in B) a 3' unsaturated aldehydic  $\alpha$ ,  $\beta$ , 4-hydroxy-2-pentanal (3'-dRP) end.

### Short or long patch BER

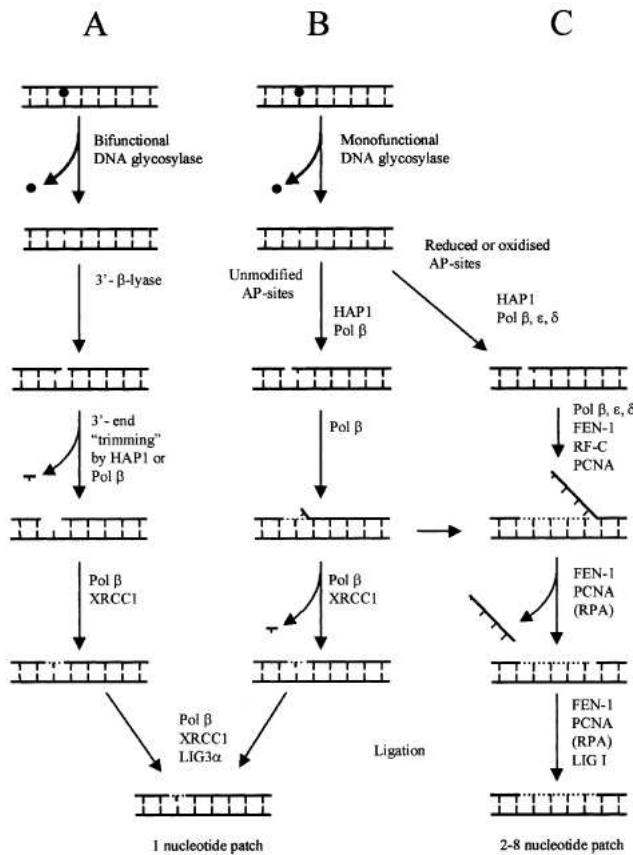
In eukaryotes the 5'-dRP is usually released by the 5'-dRPase activity of DNA polymerase  $\beta$  (Pol $\beta$ ) (RecJ in *E. coli*) and the 1 nucleotide gap is filled by the DNA polymerase activity of Pol $\beta$ . The remaining nick is sealed by the XRCC1-DNA ligase III $\alpha$  (LigIII $\alpha$ ) complex. However when the DNA glycosylase possesses an AP lyase activity as well, a 3'-dRP end is generated that requires the action of a 3' phosphodiesterase to yield a 3'-OH group, allowing Pol $\beta$  to fill the gap and the LigIII $\alpha$  complex to seal the nick. Both of these processes are known as short patch repair (Krokan, Nilsen *et al.* 2000) (Boiteux and Guillet 2004).

However in some situations when the dRP lyase activity of Pol $\beta$  is unable to release the 5'-dRP end, the 5'-flap endonuclease (Fen1) is recruited resulting in the removal of a longer patch (2-8 nucleotides) of DNA bases. Here the clamp loader RFC and PCNA are required to assist DNA polymerases  $\delta$  and  $\epsilon$  in DNA synthesis to fill the



gap, as seen in DNA replication. DNA Ligase I completes the long patch repair by sealing the nick (Krokan, Nilsen *et al.* 2000) (Boiteux and Guillet 2004).

**Figure 1.19**  
**Alternative BER**  
**pathways in**  
**mammalian cells**



A and B short patch BER repair, A using bifunctional DNA glycosylase.

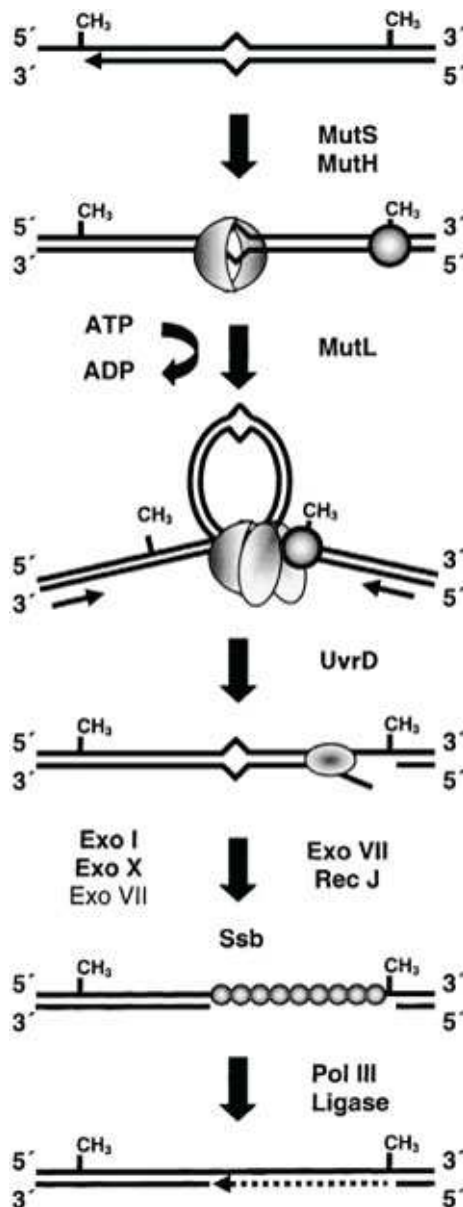
C Processing oxidised or reduced AP sites resulting via the long patch BER repair pathway. Taken from (Krokan, Nilsen *et al.* 2000)

#### 1.4.6 Mismatch repair

Mismatch repair (MMR) is required for the removal of misincorporated or unpaired bases that have arisen from DNA replication errors, spontaneous or induced base modifications and DNA recombination that would otherwise result in mutation and genomic instability (Marti, Kunz *et al.* 2002) (Yang 2000). Defects in the MMR machinery have been shown to manifest in colorectal cancers. The mismatch repair system is able to detect wrongly incorporated bases amongst normal base pairs as well as identifying eight types of mismatches, including loop structures caused by deletion or insertion in one strand. The mismatch repair system only acts on the newly synthesised daughter strand enabling the removal of the misincorporated base not the correct original base (Yang 2000).

#### **1.4.7 Mismatch repair in *E. coli***

In *E. coli* mismatch repair is carried out by three proteins MutS, MutL and MutH, which make up the MTHLS system. MutS, a homodimer, is responsible for recognizing a mismatched base pair as well as insertions or deletions of 1-4 nucleotides. Furthermore the mismatch repair system is required to distinguish between parent and daughter DNA strands and does this through recognition of methylation of the parental strand by MutH (Yang 2000). Within a few minutes after synthesis of the nascent strand during DNA replication, the deoxyadenosine methyltransferase (Dam) methylates the sequence 5'-GATC-3', converting adenine to 6-methyladenine. MutH is a GATC sequence and methylation-specific endonuclease that makes a 5' incision at the closest unmethylated GATC to the misincorporated base in hemimethylated DNA, thereby only incising the daughter strand. The role of MutL is to facilitate interaction between MutS and MutH (Yang 2000). Once the mismatch is identified and bound by MutS, hemimethylated sites are bound by MutH. Binding of ATP to MutS facilitates the formation of the DNA-MutS-MutL complex, and coupled with ATP hydrolysis by MutL, activates the endonuclease MutH. MutH incises the new strand at the 5' of the misincorporated base and the helicase UvrD is then recruited, through interaction with MutL, to unwind the duplex DNA from the nick to the misincorporated base. The single-stranded DNA is then degraded by 5'-3' or 3'-5' exonucleases, for example exonuclease VII or exonuclease I respectively, whilst the parental ssDNA is protected by SSB. The resulting gap is then filled by DNA polymerase III, which requires the presence of SSB, and the remaining nick sealed by DNA ligase (Yang, Junop *et al.* 2000).



**Figure 1.20 Mismatch repair in *E. coli***

Mismatches are recognized by MutS and hemimethylated *dam* sites are bound by MutH. MutL binds via ATP hydrolysis to form a MutHLS heteroduplex at the mismatch site. MutH activation leads to MutH incising on the newly synthesised and transiently unmethylated strand. UvrD helicase unwinds dsDNA and the resulting ssDNA containing the mismatch is degraded by nucleases 5'-3' exonucleases Exo I, X, VII or 3'-5' exonucleases. The ssDNA template strand is bound by SSB to protect it from nuclease degradation. The resulting gap is filled by DNA polymerase III and the nick sealed by DNA ligase. Taken from (Marti, Kunz *et al.* 2002)

#### **1.4.8 Mismatch repair in eukaryotes**

The eukaryotic mismatch repair (MMR) system does not use methylation as a means to discriminate between the template and nascent strand, therefore there is no MutH homologue in eukaryotes. Furthermore the MutS and MutL homologues, denominated MSH and MLH respectively, are present in different heterodimeric complexes. Different MLH/MSH heterodimers deal with distinct mismatches and DNA loops formed during DNA replication or recombination, both MutS $\alpha$  and MutS $\beta$  have a slight redundancy with each other in their recognition of insertion/deletion loops (IDLs). IDLs are caused by a differing number of micro-satellite repeat units on each DNA strand, which occurs in DNA replication when the primer and template strands

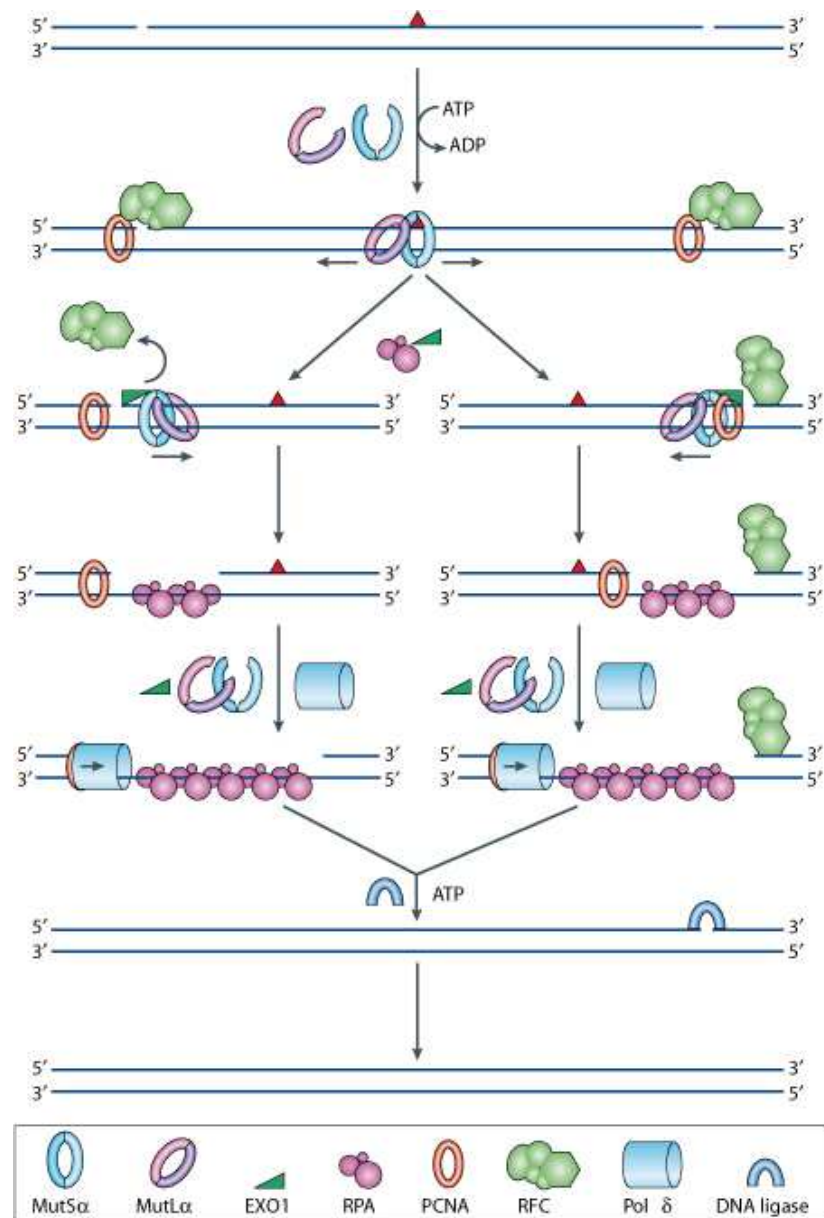
of DNA have dissociated and reannealed (Jiricny 2006). An additional MutL homolog has been identified in eukaryotes, post-meiosis segregation 1 (PMS1) in yeast (denominated PMS2 in humans), both of which form heterodimers with MLH1 in both yeast and humans (Marti, Kunz *et al.* 2002).

Complex	Component	Function
MutS $\alpha$	MSH2, MSH6	Recognition of base-base mismatches and small IDLs
MutS $\beta$	MSH2, MSH3	Recognition of IDLs
MutL $\alpha$	MLH1, PMS2	Forms a ternary complex with mismatch DNA and MutS $\alpha$ ; increases discrimination between heteroduplexes and homoduplexes; also functions in meiotic recombination
MutL $\beta$	MLH1, PMS1	Unknown
MutL $\gamma$	MLH1, MLH3	Primary function in meiotic recombination; backup for MutL $\alpha$ in the repair of base-base mismatches and small IDLs

**Table 1.2 Human MutS and MutL homologue complexes**

MLH and MSH are MutL and MutS homologues, respectively. The subunits that make up different MutS and MutL complexes are shown with their specific roles in mismatch repair including the removal of insertion or deletion loops (IDLs) arising from micro-satellite repeat units. Post-meiotic segregation protein (PMS). Adapted from (Jiricny 2006).

Discrimination between the parent and daughter strands arises from interaction of both MSH and MLH with the proliferating cell nuclear antigen (PCNA), a  $\beta$  sliding clamp and processivity factor for the DNA polymerase. PCNA has a distinct orientation on the DNA that allows the MSH and MLH proteins to identify the daughter strand (Marti, Kunz *et al.* 2002). PCNA plays a fundamental role in eukaryotic MMR due to a strong interaction with the complexes MSH2-MSH6 and MSH2-MSH3 (Clark, Valle *et al.* 2000), modulation of 3' and 5' excision (Guo, Presnell *et al.* 2004) and interaction with DNA polymerase  $\delta$  facilitating the gap-filling stage of MMR (Gu, Hong *et al.* 1998).



**Figure 1.21 Human mismatch repair pathway**

The mismatch denoted by a red triangle is bound by either MutSα or MutSβ that loads MutLα. ATP hydrolysis facilitates a conformational change of the clamps that leads to their migration. Clamps that migrate upstream to the RFC bound at the 5' end of the DNA break, displace RFC and loads the nuclease EXO1. EXO1 degrades the DNA strand in a 5'-3' direction and the resulting ssDNA is bound and stabilized by RPA. Once the mismatch has been degraded, MutSα stimulation is replaced by MutLα inhibition. Pol and PCNA load at the 3' end and commence DNA synthesis to fill the gap, and the remaining nick is sealed by DNA ligase (a). Alternatively the migration of the clamps downstream leads to the encounter of the 3' end bound by PCNA. EXO1 is recruited and degradation begins, involving several reloading events of EXO1 until the mismatch has been removed. RPA bound to ssDNA and MutLα inhibit any further degradation, the resulting gap is filled by Polδ and DNA ligase seals the nick. Taken from (Jiricny 2006).

#### **1.4.9 Homologous recombination**

Homologous recombination can repair a DSB as well as restarting DNA replication from a stalled and detached replication fork, and is a relatively error-free repair process (Dillingham and Kowalczykowski 2008). Homologous recombination is also required during conjugation and meiosis, where homologous chromosomes undergo crossing over via homologous recombination to generate a new combination of alleles (Schwartz and Heyer 2011). Non-homologous end joining is capable of repairing DSBs but is a highly error prone process, as discussed later (Pitcher, Brissett *et al.* 2007).

Homologous recombination can be divided into several stages. Presynaptic processing, where the DNA ends are resected to generate 3' overhangs, is performed by a variety of enzymes depending on the circumstances and the organism. Both the mode and enzymes involved are not conserved from bacteria to eukaryotes, and are yet to be determined in archaea. Synaptic formation, where recombinase monomers polymerise onto the resulting ssDNA, identifies homologous intact dsDNA and catalyses strand exchange resulting in the formation of branched DNA structures. This is the central step of homologous recombination and is performed by the RecA family of recombinases, it is highly conserved across the three domains of life. When both strands of the DNA are damaged, for example from a replicated adduct, chromosome break or cross-link, the damaged region must be resynthesized using intact homologous donor DNA as a template (Dillingham and Kowalczykowski 2008). The identification of homologous donor DNA and subsequent strand exchange is performed by recombinase RecA in *E. coli* and the homologs RAD51 in eukaryotes and RadA in archaea. The recombinase monomers polymerise onto ssDNA forming a helical nucleofilament and act as a scaffold binding ssDNA and dsDNA in an extended, unwound conformation that promotes base pairing (Bell 2005).

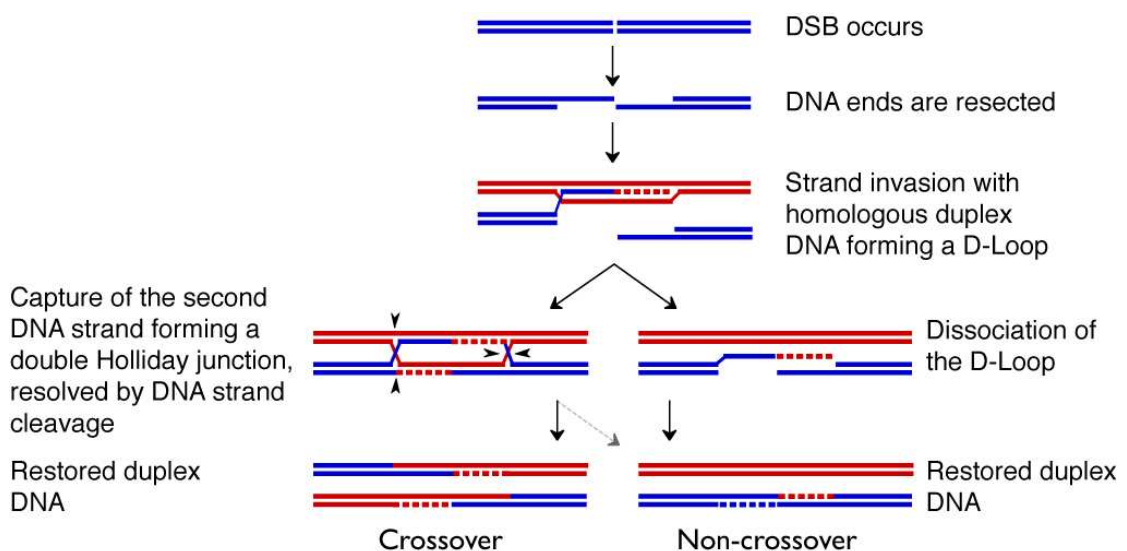
Once bound to the ssDNA, the dynamic recombinase nucleoprotein filament scans neighbouring intact dsDNA for homologous dsDNA regions. Once homologous intact dsDNA is found the recombinase catalyses the switch of base-paired partners, resulting in the incoming ssDNA molecule being base paired with the complementary strand from the intact dsDNA molecule (Holthausen, Wyman *et al.* 2010).



**Figure 1.22 Formation of a replication fork using homologous recombination**

The formation of D-loop followed by branch migration leads to a Holliday junction. Cleavage of the Holliday junction can result in the formation of a replication fork.

The formation of a D-loop and subsequent migration results in a four-armed DNA junction intermediate called a Holliday junction (Schwartz and Heyer 2011) shown in Figure 1.23, which was first identified by Robin Holliday as a means by which crossovers occur during DNA repair (Holliday 1964). If the second ssDNA strand end is captured a double Holliday junction will be formed.



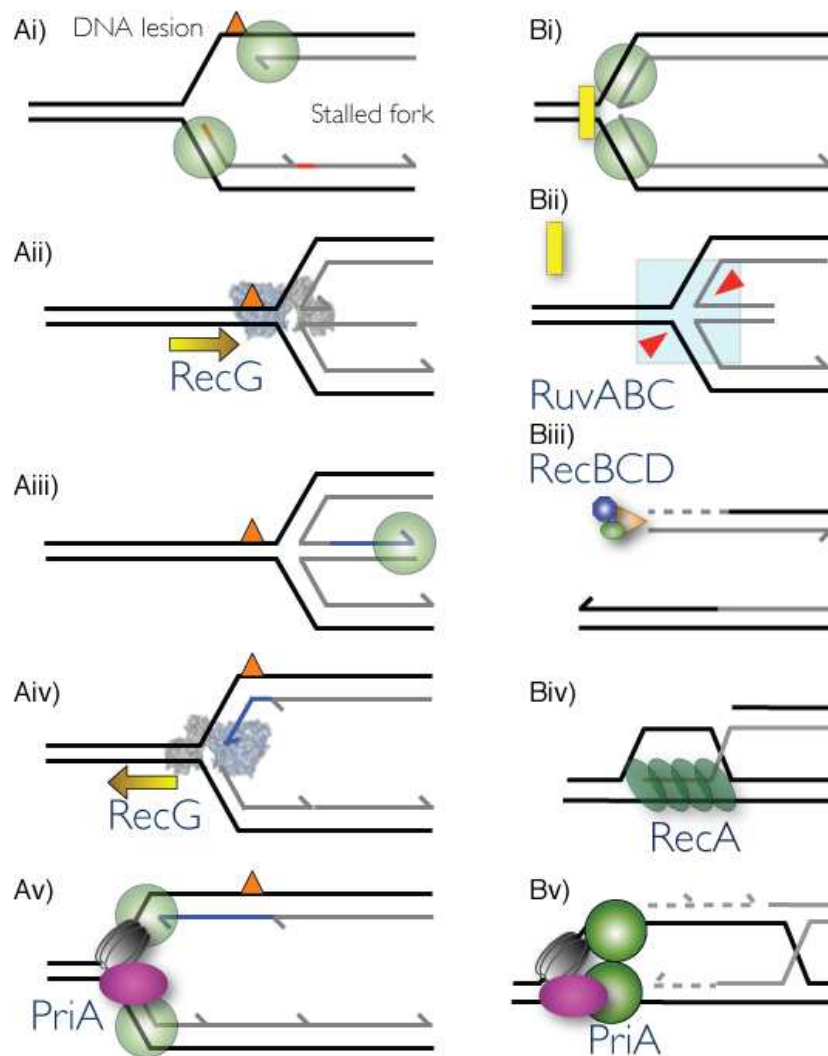
**Figure 1.23 Homologous recombination**

Homologous recombination model resulting in a cross over between the two homologous DNA chromatids, however if second end capture does not occur and the D-Loop dissociates, a non-crossover product is generated through synthesis dependent strand-annealing (SDSA).

Once DNA synthesis has been performed to repair the break using the homologous intact DNA template, the branched structures are resolved restoring the DNA to its duplex structure. (San Filippo, Sung *et al.* 2008). If only one Holliday junction has been formed, synthesis dependent annealing (SDA) can occur. Homologous

recombination can also be used to form a replication fork, as shown in Figure 1.22. Recombination proteins can also be used to restart a replication fork but strand exchange and the formation of a D-loop is not always required, as shown in Figure 1.23.



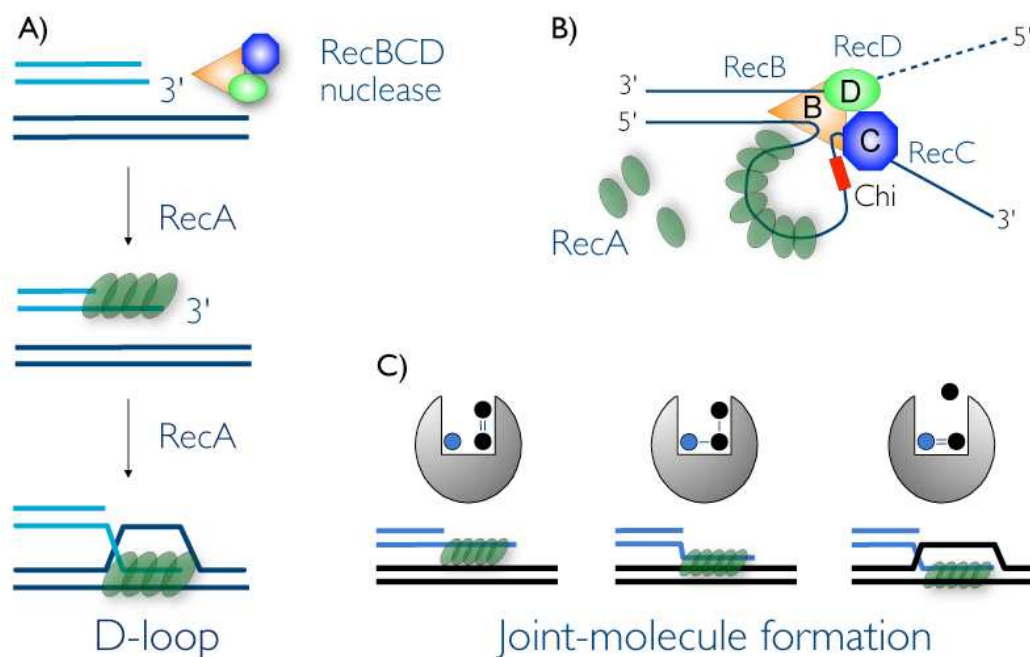


**Figure 1.24 Restarting a stalled replication fork using homologous recombination proteins**

Two different mechanisms can be employed to restart a stalled replication fork, (A) by manipulating the replication fork to use the newly synthesised strand as a template to bypass the DNA lesion and (B) using cleavage of a DNA replication fork and strand exchange to remove a DNA lesion and reform a replication fork. *E. coli* proteins depicted, DNA polymerase shown by green circle. DNA lesion causes replication fork to stall (Ai) and is pushed back past the DNA lesion using a DNA helicase (Aii). A DNA polymerase can fill in the ssDNA end (blue arrow line) or a nuclease can degrade the ssDNA (Aiii). The helicase then unwinds duplex DNA to restore the fork to its original position (Aiv) and DNA synthesis continues (Av). Alternatively stalled replication fork is pushed back and incised by RuvABC, removing the DNA lesion (Bii). The DNA ends are then processed to generate a 3'-overhang (Biii), to which RecA polymerises onto and catalyses strand exchange to form a new replication fork (Bv).

#### 1.4.10 Homologous recombination in *E. coli*

The presynaptic processing of DNA ends to generate the 3' ssDNA overhang, (Holthausen, Wyman *et al.* 2010) is performed by the RecBCD complex. RecBCD achieves this through coordination of enzymatic activities; RecB is a 3'-5' helicase and multifunctional nuclease, RecC recognizes the Chi sequence and RecD is 5'-3' helicase (Singleton, Dillingham *et al.* 2004). When the RecBCD complex reaches the Chi sequence (5'-GCTGGTGG-3') during resection, the 5'-3', nuclease activity of RecBCD is significantly decreased whilst the 3'-5' nuclease activity is increased, thereby generating a 3' overhang. RecA is loaded onto the ssDNA 3' tail during RecBCD resection. However until RecA polymerises onto the ssDNA, the latter is prevented from forming secondary structures and protected from nuclease degradation by SSB (Yeeles and Dillingham 2010).

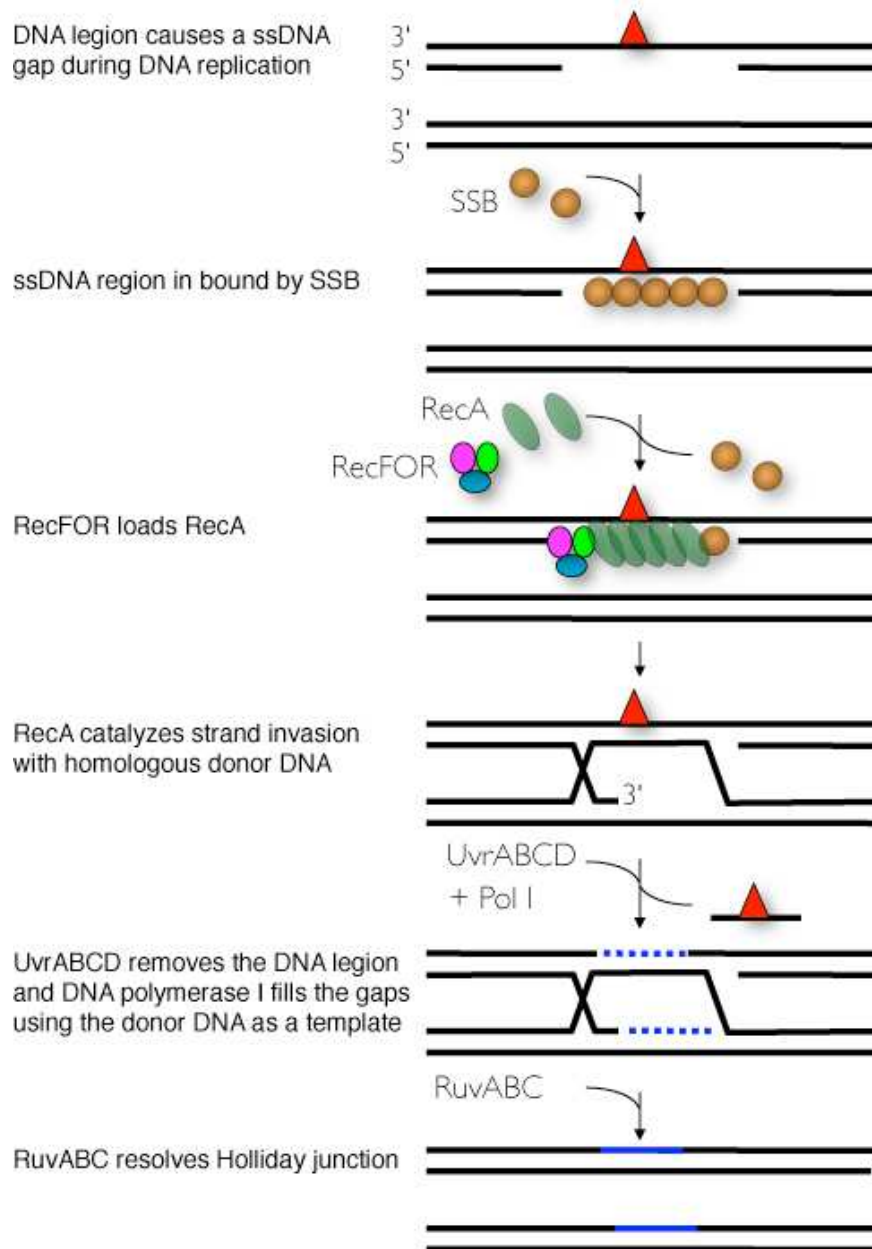


**Figure 1.25 Activity of RecBCD and loading of RecA**

A) RecBCD resects DNA ends generating a 3' overhang onto which the RecA recombinase polymerises, scans duplex DNA for homologous regions and catalyses strand exchange forming a D-loop. B) RecBCD directly loads RecA whilst processing the DNA ends. C) RecA transfers base-pairing from homologous donor duplex DNA (black circles) to damaged invading DNA strand in the formation of a joint-molecule.

Although RecBCD is the primary complex for the processing of the DNA ends, the helicase RecQ and exonuclease RecJ are able to fulfil a similar role. The proteins

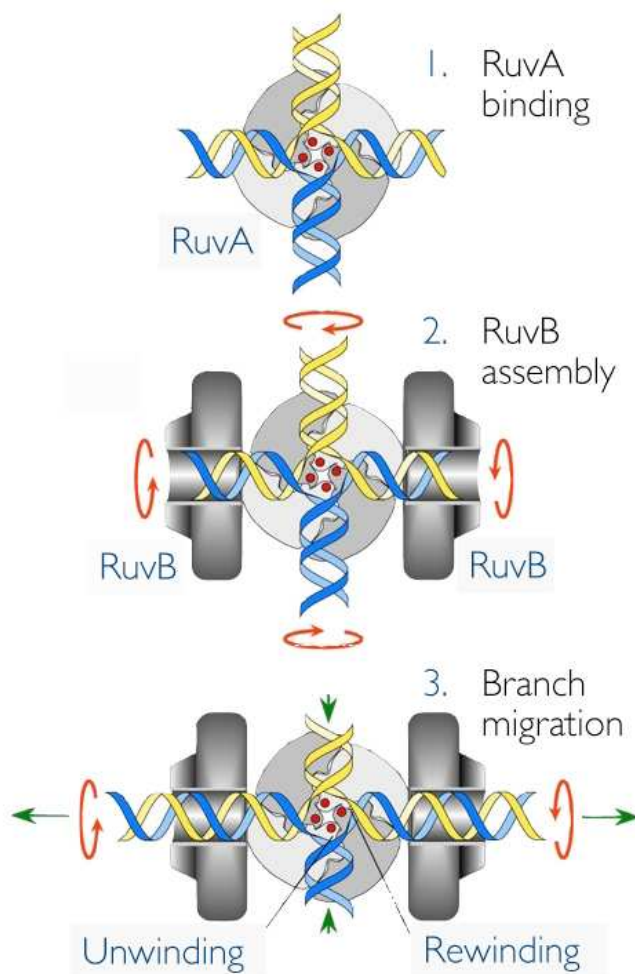
RecF, RecO and RecR, form both RecOR and RecFR complexes. The RecOR complex is required for the loading of RecA onto the ssDNA due to the presence of SSB. The RecFR complex prevents the RecA filamenting onto dsDNA where it would be inactive (Yeeles and Dillingham 2010). The RecFOR system is able to load RecA for the repair of ssDNA gaps that have arisen due to the presence of a DNA lesion that has not been removed during DNA replication, as shown in Figure 1.26.



**Figure 1.26 ssDNA gap repair by homologous recombination**

RecFOR loads the RecA recombinase to facilitate repair of a ssDNA gap using homologous recombination. UvrABCD catalyzes the removal of the DNA lesion and DNA polymerase I (Pol I) synthesis DNA to fill the remaining DNA gaps.

The nucleoprotein filament will then scan intact donor DNA for the homologous region and catalyse strand exchange between the two DNA molecules, forming a joint molecule as discussed before. Optionally, the second 3' end of the damaged incoming DNA can also be engaged. The DNA polymerase extends the invading 3' end(s) of the damaged DNA using the homologous donor DNA as a template, forming a Holliday junction, as shown in Figure 1.26. Branch migration is mediated by RuvA, a Holliday junction binding protein, and RuvB, a ring helicase, of the RuvABC complex.

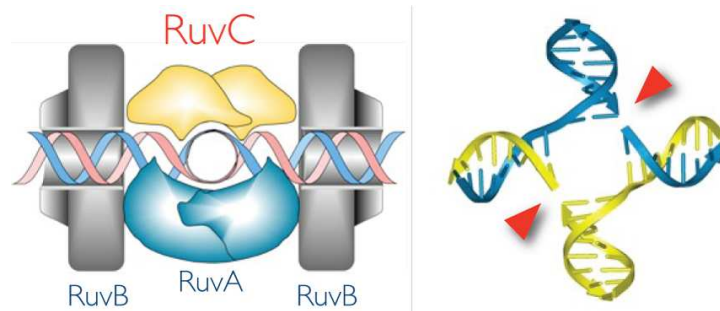


**Figure 1.27 Resolution of a Holliday junction by RuvABC.**

1. RuvA recognizes and binds the Holliday junction. 2. RuvB is then loaded onto the Holliday junction either side of RuvA. 3. Branch migration is facilitated by the RuvB and DNA incisions are performed by RuvC.

RuvA and B unwind the dsDNA to extend the ssDNA template region available to the DNA polymerase (Donaldson, Courcelle *et al.* 2006). Resolution is performed by the structure-specific endonuclease RuvC of the RuvABC complex, producing either crossover or non-crossover ligatable products (Holthausen, Wyman *et al.* 2010).

However RecG, a branch specific helicase, is able to unwind structures including Holliday junctions and D-loops into duplex DNA. The degree of sensitivity to DNA damaging agents of the  $\Delta recG$  and *ruv* mutations compared with single  $\Delta recG$  suggests overlapping functions between RuvABC and RecG in homologous recombination (Rudolph, Upton *et al.* 2010).

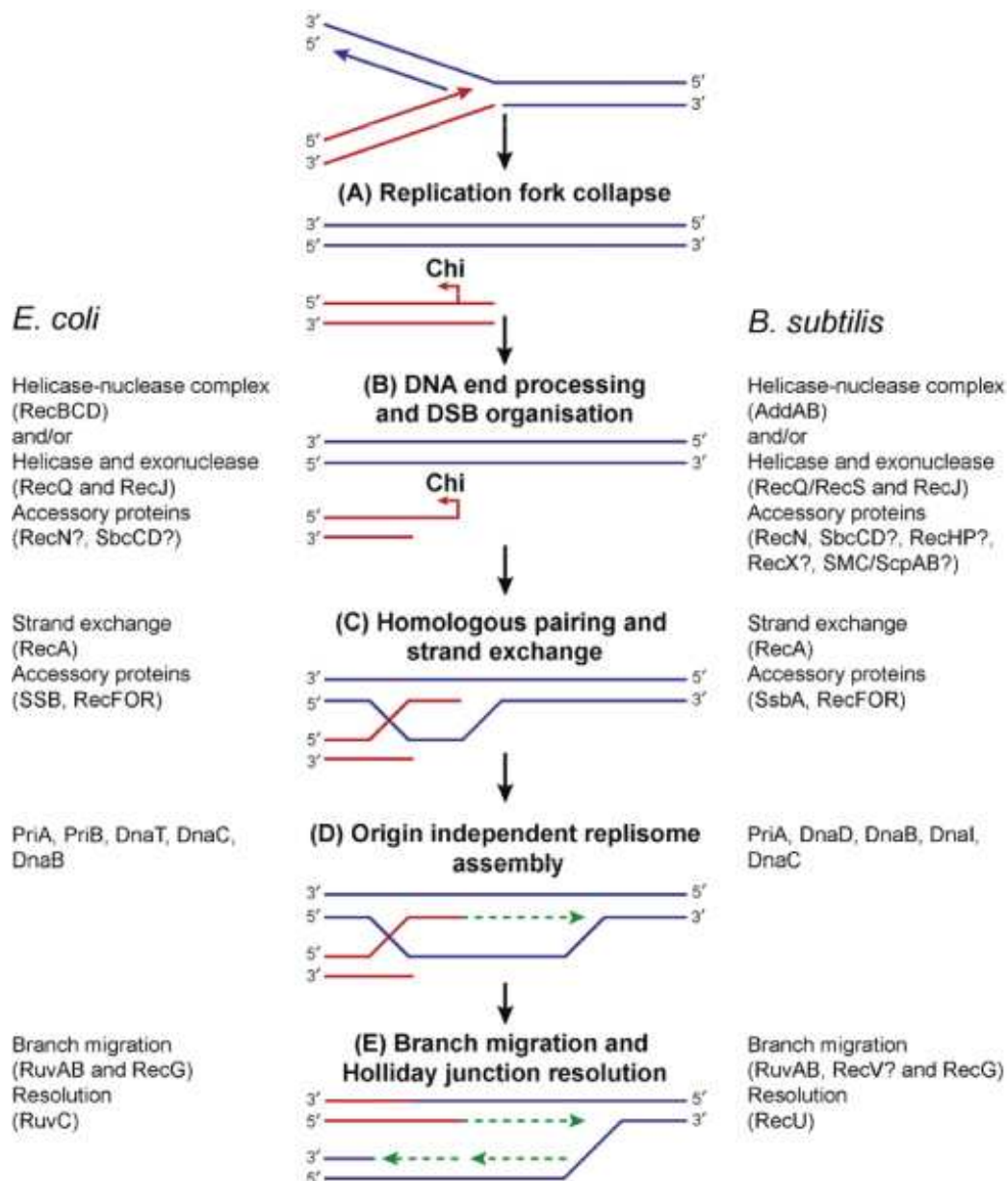


**Figure 1.28 RuvABC resolves a Holliday junction**

The structure of RuvABC bound to DNA in a Holliday junction is shown on the left. Where RuvABC incises a Holliday junction is shown on the right.

#### **1.4.11 Homologous recombination in *Bacillus subtilis***

The helicase-nuclease complex responsible for DNA end processing in *B. subtilis* is the AddAB complex. Although the AddAB complex is functionally similar to the *E. coli* RecBCD complex, it is structurally different, the AddAB complex being a heterodimer of AddA and AddB subunits. The AddA subunit contains a helicase and nuclease domain resembling RecB and cleaves the 3'-5' strand. The AddB subunit possesses a RecB-family nuclease domain that cleaves 5'-3'. Both nuclease activities of AddA and AddB are equal until the Chi sequence is recognized. Chi recognition results in the attenuation of the 3'-5' nuclease activity, generating a 3'-end overhang by the continued 5'-3' nuclease activity. Although there are subtle differences between the *E. coli* and *B. subtilis*, including the Chi being a pentameric 5'-AGCGG-3' sequence, the mechanistic process of homologous recombination is similar to that used in *E. coli*, as shown in Figure 1.29 (Yeeles and Dillingham 2010).



**Figure 1.29 Homologous recombination in *E. coli* and *B. subtilis*.**

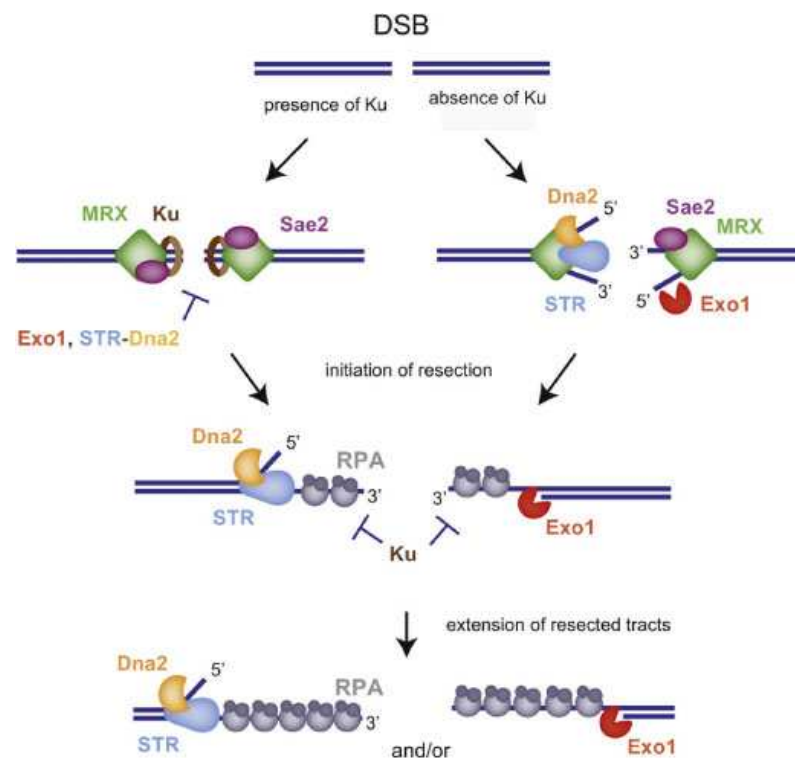
The HR pathway to rescue collapsed replication forks, which is the same for DSB repair is shown for both *E. coli* and *B. subtilis*. Proteins whose function in the process is yet to be determined are indicated with a question mark. Taken from (Yeeles and Dillingham 2010).

#### **1.4.12 Homologous recombination in eukaryotes**

As stated earlier, the key step in homologous recombination is the formation of the recombinase nucleoprotein filament on ssDNA. The eukaryotic recombinase Rad51



achieves this through interaction with recombinase mediator BRCA2 (Breast Cancer 2 susceptibility protein) in humans and Rad52 in yeast. Before this can occur the damaged DNA ends are processed to generate a 3'-end ssDNA overhang that Rad51 can polymerise onto. The Mre11-Rad50-Nbs1 (MRN) complex in human and the Mre11-Rad50-Xrs2 (MRX) complex in *S. cerevisiae* play a central role in the initiation of DNA end resection by recognizing the dsDNA break, and removing a short oligonucleotide. This is followed by the recruitment of a variety of nucleases including Exo1 and/or Dna2 to the DNA ends. The Sgs1-Top3-Rmi1 (STR) complex is also recruited to the DSB in yeast, and is responsible for the unwinding of the DNA at the DSB to allow nuclease degradation (Mimitou and Symington 2011).



**Figure 1.30 DSB end resection**

Taken from (Mimitou and Symington 2011). The Mre11-Rad50-Xrs2 (MRX complex) binds to a DSB and facilitates DSB recognition by checkpoint machinery, tethering the DSB ends and end-processing for repair by homologous recombination. MRX-Sae2 catalyse and initial resection of a short oligonucleotide fragment from the 5' ends. Followed by extensive degradation by two parallel pathways requiring Sgs1-Top3-Rmi1 (STR) with Dna2 (left-hand model) or Exo1 (right-hand model). MRX can function as a scaffold for resection machinery to bind to (right-hand model). The loss of Tku70-Yku80 (Ku), which can recognize a DSB and inhibit the loading of Exo1 and Sgs1 by MRX, shows overlap between the two pathways.

The ssDNA produced by nuclease activity is bound by RPA to ensure both that the ssDNA does not form secondary structures, or is degraded by nucleases. The elimination of secondary structures stimulates homologous recombination but at the same time the presence of RPA bound to the ssDNA inhibits the formation of the recombinase Rad51, nucleoprotein filament (Shinohara, Shinohara *et al.* 1998). In order to displace RPA and load the RecA homolog, Rad51, the assistance of the recombination mediators Rad52 in yeast, and BRCA2 in humans is required. Both Rad52 and BRCA2 target Rad51 to the ssDNA at a DSB, and both BRCA2 and Rad52 bind the ssDNA directly and interact with RPA in order to load Rad51 (San Filippo, Sung *et al.* 2008). BRCA2 possesses a highly conserved BRC repeat domain, consisting of eight copies of a 30-40 amino acid sequence, which interacts directly with Rad51 as well as OB-folds that are used to bind ssDNA (Forget and Kowalczykowski 2010). The Rad51 nucleoprotein filament then binds an intact homologous duplex DNA, scans for a region of homology and catalyses strand exchange. Rad52 is able to assist Rad51 in the search for homologous DNA.

After the invasion of the first damaged DNA strand into the homologous duplex DNA, catalysed by Rad51, a D-loop is formed and at this stage HR is not committed, as the D-loop can dissociate through the action of the translocase Rad54 (Bugreev, Hanaoka *et al.* 2007). If HR continues, DNA synthesis is followed by dissociation of the D-loop and leads to synthesis dependent strand annealing, as shown in Figure 1.23 in the strand displacement and strand annealing section (San Filippo, Sung *et al.* 2008).

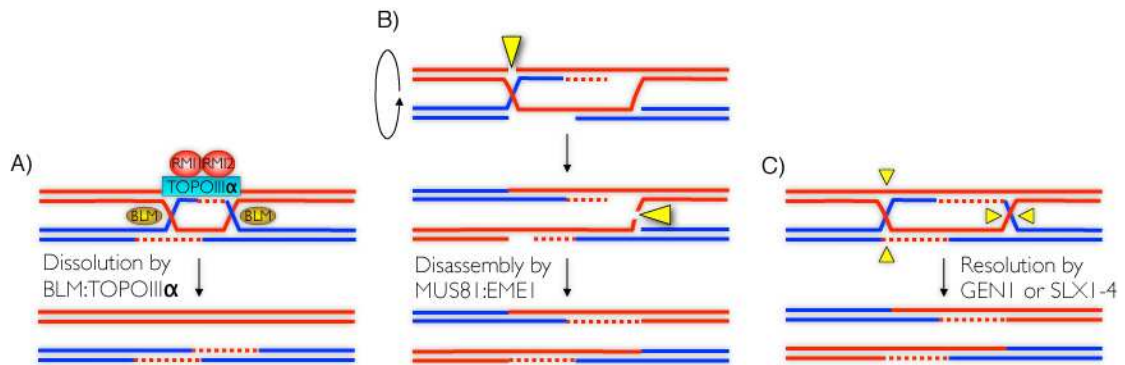
Optionally, the second ssDNA strand is then captured by Rad51 and bound to the homologous donor DNA to form a second Holliday junction, shown in Figure 1.23. The protein complex HOP2-MND1 plays a crucial role in stabilising the Rad51 nucleofilament as well as assisting the capture of duplex donor DNA to enhance the formation of branched synaptic complexes (San Filippo, Sung *et al.* 2008).

In order for the DNA polymerase to extend the 3' end at the D-loop, the translocase Rad54 is required to dissociate the Rad51 nucleoprotein filament from the 3' end. Rad54 also stimulates the search for homologous donor DNA, chromatin remodelling, formation of the branched D-loop formation resulting from the invasion of the first strand of DNA into the duplex donor DNA, and the postsynaptic removal of Rad51



from the restored dsDNA. As synthesis of DNA progresses, Rad54 and other proteins including FANCM can mediate the migration of the branched DNA. (Boddy, Gaillard *et al.* 2001). The FANCM protein, one of thirteen Fanconi anemia proteins, possesses a N-terminal helicase domain and a C-terminal nuclease domain. FANCM specifically binds branched DNA structures including Holliday junctions and has been shown *in vitro* to act as a translocase facilitating branch migration (Gari, Decaillet *et al.* 2008) and displacing the third strand of a triple helix (Meetei, Medhurst *et al.* 2005).

The resolution of a Holliday junction in eukaryotes is not limited to a single unified model as seen in bacteria; the plasticity of individual nuclease uses in distinct functional contexts varies from species to species. Holliday junctions can be resolved by an endonuclease and also separated by dissolution, shown in Figure 1.31. In eukaryotes, the XPF-family endonuclease Mus81-Mms4/EME1 can function in the dissolution step, as can the STR complex, in a more RecG-like manner by unwinding and temporarily nicking the DNA. The eukaryotic Yen1/GEN1 Rad2/XPG endonuclease was identified as a Holliday junction resolvase, however Yen1/GEN1 mutants in three different major organisms including yeast and humans, display no major phenotypes in cell growth or DNA repair. Therefore, the only endonuclease to display the biochemical hallmarks of the bacterial RuvC resolvase appears not to be essential, raising the issue that not all Holliday junction resolvase candidates have been identified (Schwartz and Heyer 2011).



**Figure 1.31. Possible Holliday junction resolution mechanisms in eukaryotes.**

A) Dissolution of a Holliday junction by RecQ helicase BLM and topoisomerase III $\alpha$ , involving the unwinding of the Holliday junction and temporary nicking of DNA. B) Single consecutive incisions by Mus81-EmeI disassemble the Holliday junctions. C) Resolution of a Holliday junction in a RecG manner, as seen in *E. coli* by Gen1 or SLX1-4.

HR repair of DSB in yeast is greatly similar to that in humans, see Table 1.3 for protein homologs. However, there are some differences, for example neither *S. cerevisiae* and *S. pombe* possess a BRCA2-like protein (San Filippo, Sung *et al.* 2008).

Human	<i>S. cerevisiae</i>	Biochemical function	Additional features
<b>Proteins that function with Rad51</b>			
MRN complex: Mre11-Rad50-Nbs1	MRN complex: Mre11-Rad50-Xrs2	DNA binding Nuclease activities	Involved in DNA damage checkpoints. Associated with DSB end resection
BRCA2	None	ssDNA binding Recombination mediator	Interacts with RPA, Rad51, Dmc1, PALB2, DSS1. Member of Fanconi anemia group
Rad52*	Rad52	ssDNA binding and annealing Recombination mediator	Interacts with Rad51 and RPA
None found to date	Rad59	ssDNA binding and annealing	Interacts with Rad52 Homology to Rad52
Rad54 Rad54B	Rad54 Rdh54	ATP-dependent dsDNA translocase. Induces superhelical stress in dsDNA. Stimulates D-loop interaction	Member of the Swi2/Snf2 protein family. Chromatin remodeler. Interacts with Rad51. Yeast proteins remove Rad51 from dsDNA

Rad51B-Rad51C Rad51D-XRCC2 Rad51C-XRCC3	Rad55-Rad57	ssDNA binding Recombination mediator activity	Rad51B-Rad51C and Rad51D-XRCC2 form a tetrameric complex. Rad51C associates with Holliday junction resolvase activity
Hop2-Mnd1	Hop2-Mnd1	Stimulates the D-loop reaction, Stabilizes the presynaptic filament. Promotes duplex capture	Interacts with Dmc1 and Rad51
<b>Proteins that function with Dmc1</b>			
Hop1-Mnd2	Hop2-Mnd1	Stimulates the D-loop reaction, Stabilizes the presynaptic filament. Promotes duplex capture	Interacts with Dmc1 and Rad51
None found to date	Mei5-Sae3	Predicted recombination mediator activity	Interacts with Dmc1. Likely functional equivalent of <i>S. pombe</i> Sfr1-SwiS
Rad54B	Rdh54	Stimulates D-loop reaction	Interacts with Dmc1 and Rad51

**Table 1.3 Eukaryotic homologous recombination factors**

Eukaryotic homologous recombination factors for yeast and humans. \*Recombination mediator activity only found in yeast Rad52. Adapted from (San Filippo, Sung *et al.* 2008).

#### **1.4.13 Meiotic recombination**

Chiasmata, are the product of crossovers between two homologous non-sister chromatids, and result from homologous recombination during meiosis, which occurs in human germ cells. This generates new combinations of alleles that could lead to a ‘fitter’ progeny. Failure to generate chiasmata leads to an abnormal number of chromosomes, known as aneuploidy, and can manifest in a variety of genetic disorders (Hassold, Hall *et al.* 2007). The meiosis specific recombinase DMC1 is present in almost all eukaryotes, it is structurally related to RecA and Rad51 recombinases. Dmc1 is required for both meiotic recombination and chromosome segregation, it polymerises onto ssDNA to form helical nucleoprotein filaments and catalyses homologous DNA pairing, as seen with RecA and Rad51 (San Filippo, Sung *et al.* 2008).

#### **1.4.14 Homologous recombination in archaea**

The Mre11 and Rad50 proteins that perform the initial steps of end resection in eukaryotes are also conserved in archaea. The Mre11-Rad50 complex has been shown to be involved in HR repair of DSBs in *Haloferax volcanii* (Delmas, Shunburne *et al.* 2009). In thermophiles genes encoding the proteins HerA, a hexameric helicase, and the nuclease NurA have been found in an operon with *mre11* and *rad50*. Whilst all four proteins have been shown to catalyse the 3'-end resection of HR in *Pyrococcus furiosus* *in vitro*, identification of the exact process and the enzymes involved in archaea is yet to be determined (White 2011). The archaeal recombinase RadA, a RecA/Rad51 homolog, possesses a 40% sequence similarity to Rad51 and only a 20% sequence similarity to RecA (Seitz, Brockman *et al.* 1998). In addition to a high conservation of the N-terminal domain between RadA and Rad51 (Guy, Haldenby *et al.* 2006). Furthermore the Rad51/RadA polymerisation motif comprises a highly conserved phenylalanine residue motif that inserts into the hydrophobic pocket of the adjacent monomer, acting as a ball-and-socket polymerisation mechanism (Shin, Pellegrini *et al.* 2003; Haldenby, White *et al.* 2009).

Once the 3'-end overhangs have been generated they are bound by the ssDNA-binding protein RPA to protect the ssDNA from forming secondary structures, or from being degraded by nucleases (Seitz, Brockman *et al.* 1998). In order for RadA to polymerise onto the ssDNA to form a nucleoprotein filament, the recombination mediator and RadA paralogue RadB is required. RadB is only found in euryarchaea, and is unable to catalyse strand exchange and consequently is not considered to be a recombinase (Komori, Miyata *et al.* 2000). Indeed, RadB has been shown to act in the same repair pathway as RadA in *H. volcanii*, interacting with RadA, Holliday junction resolvase (Hjc) and DNA polymerase II (Guy, Haldenby *et al.* 2006) (Komori, Miyata *et al.* 2000) (Hayashi, Morikawa *et al.* 1999). RadB knockouts in *Haloferax volcanii* display growth, DNA repair and DNA recombination defects similar to those of a double  $\Delta radA \Delta radB$  deletion, again suggesting they function in the same repair pathway. Mutation in the hydrophobic pocket of RadA of an alanine to a valine increases the hydrophobicity of a pocket to levels seen in crenarchaea. This suppressor mutation restores the  $\Delta radB$  strain growth, DNA repair and DNA recombination to wild-type levels (Haldenby 2007), suggesting RadB functions as a recombination mediator (Haldenby, White *et al.* 2009). RadB binds with a higher affinity to ssDNA

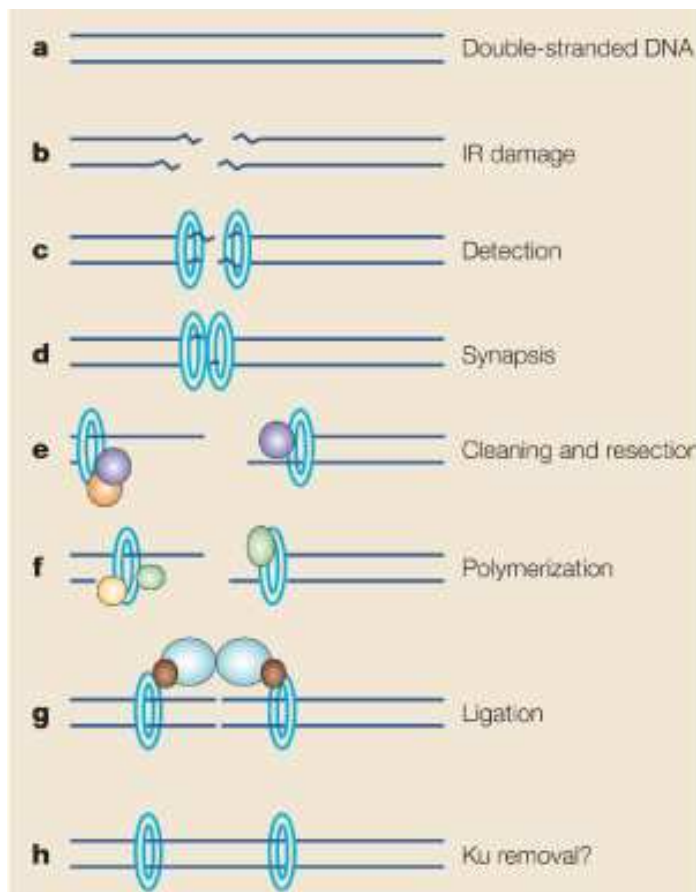
than RadA, and is thought to displace ssDNA bound RPA, facilitating the polymerisation of RadA onto the ssDNA. In the absence of RadB, the suppressor mutation is thought to stabilize and strengthen RadA:RadA monomer interaction, enabling RadA to displace RPA bound to the ssDNA. RadA, once bound to the ssDNA identifies homologous DNA in an intact sister chromatid and catalyses strand exchange as described in eukaryotes (Seitz, Brockman *et al.* 1998).

In *Pyrococcus furiosus* RadB has been shown to have a regulatory role on the activity of Hjc, perhaps ensuring cleavage of the Holliday junction at the appropriate moment (Komori, Miyata *et al.* 2000). All archaeal genomes sequenced to date encode a Hjc-like protein, however there is no identifiable sequence similarity with any other nuclease (Daiyasu, Komori *et al.* 2000). Hjc is able to cleave the Holliday junction formed during HR, resulting in two sister chromatids one of which possesses ligatable nicks (Nishino, Komori *et al.* 2001). The archaeal Xpf/Mus81/FANCM homolog Hef, (helicase-associated endonuclease fork-structure) possesses a DEAH N-terminal helicase domain, and a XPF C-terminal endonuclease domain. Hef is able to act on nicked, flapped and forked DNA and consequently can convert a Holliday junction into a forked structure. The synthetic lethality of  $\Delta hef \Delta hjc$  suggests that Hef and Hjc do not act in the same pathway. Coupled with the essentiality of the helicase and nuclease activities of Hef in a  $\Delta hjc$  background, it has been proposed that there are overlapping functions between Hef and Hjc and that Hef can act as an alternative Holliday junction resolvase (Lestini, Duan *et al.* 2010). This functional overlap between two Holliday junction resolvases is also seen in *S. cerevisiae*, between Mus81 and Yen1 in HR repair of DNA (Blanco, Matos *et al.* 2010).

#### **1.4.15 Non-homologous end joining**

Non-homologous end joining (NHEJ) is a major repair pathway of DSBs that involves the re-joining of two damaged DNA ends. It is not a homology-directed repair pathway like HR, and for this reason is both error-prone and necessary for the repair of DSBs outside S/G2 phase of the cell cycle, when there is no homologous sister chromatid available for HR. NHEJ is a flexible pathway that involves enzymes that act repetitively, in any order and independently of each other at the ends of a DSB (Lieber 2010). NHEJ is considered an error prone repair pathway unlike HR, due to

imprecise alignment of 3' ssDNA overhangs generated during NHEJ. This can result in erroneous pairing and/or deletion of base pairs that NHEJ is unable to recognize and correct (Daley, Palmbos *et al.* 2005). The basic mechanism of NHEJ is conserved from bacteria to eukaryotes, as shown in Figure 1.32, with different enzymatic homologs shown in Table 1.4. Mammals alone possess the nuclease Artemis, which functions as a 5' and 3' endonuclease and is able to open up hairpins formed by the damaged DNA ends to allow NHEJ to proceed (Lieber 2010). This function is provided in yeast by Mre11-Rad50-Xrs2 (MRX). NHEJ machinery is present in a limited number of archaea across the taxa, including Euryarchaeota and Korarchaeota, but the exact process is yet to be determined (Smith, Nair *et al.* 2011).



**Figure 1.32 The mechanism of non-homologous end joining**

Double-stranded DNA (a) is damaged resulting in a DSB (b). The DSB is detected and each DNA end bound by Ku (c). A synapse is formed to keep the two DNA ends in close proximity to each other to allow DNA repair (d). If the ends are 'dirty' nuclease resection is required to generate 3' overhangs (e) to allow resynthesis by a DNA polymerase (f). In order for the nicks to be religated they are brought into close proximity through the formation of a synapse (g). However it is not known at what stage Ku is removed from the DNA (h). Taken from (Downs and Jackson 2004).

Functional component	Bacteria	Eukaryotes	
		<i>S. cerevisiae</i>	Multicellular Eukaryotes
<b>Toolbelt protein Polymerase</b>	Ku (30-40 kDA) PolD domain of LigD	Ku 70/80 Pol4	Ku 70/80 $\mu$ and $\lambda$
<b>Nuclease</b>	?	Rad50:Mre11:Xrs2 (FEN-1)	Artemis:DNA-PKcs
<b>Kinase/ Phosphate</b>	PE domain of LigD	Tpp1 and others	PNK and others
<b>Ligase</b>	LIG domain of LigD	Nej1:Lif1:Dnl4	XLFXRCC4:DNA ligase IV

**Table 1.4 Bacterial and eukaryotic non-homologous end joining enzymes**

The enzymes involved in NHEJ for bacteria and eukaryotes. Adapted from (Lieber 2010).

#### **1.4.16 Non-homologous end joining in eukaryotes**

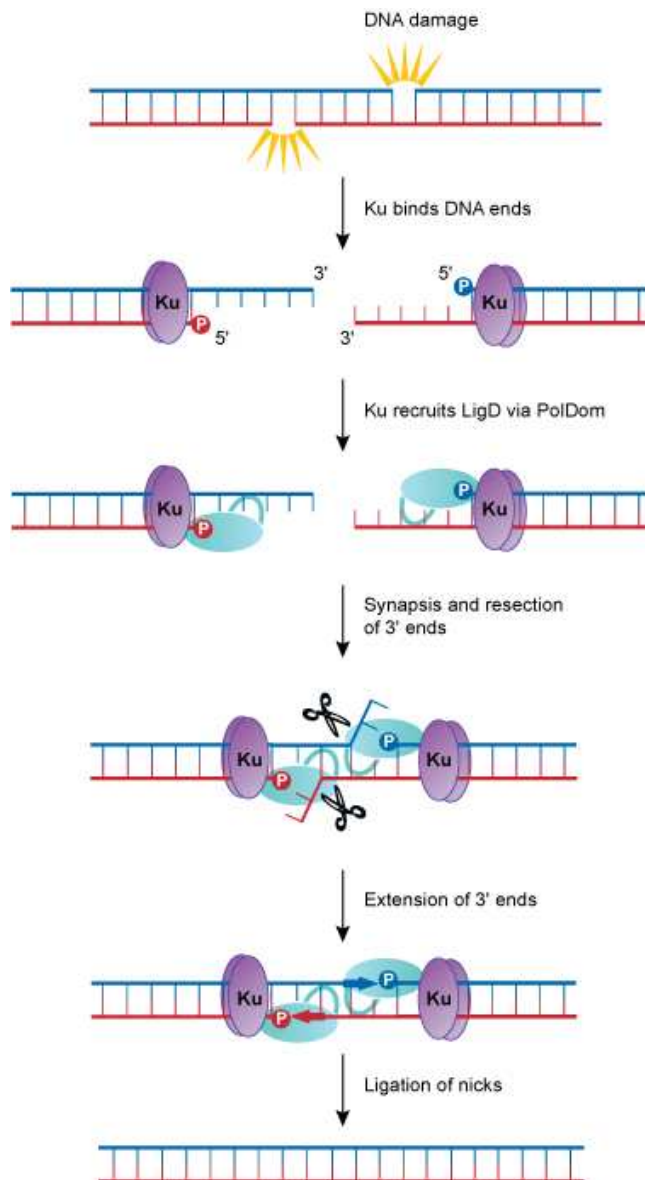
The protein Ku recognizes a DSB and binds each DNA end forming Ku:DNA end complexes. Ku is able to act as a toolbelt, both recruiting and providing a docking platform for nucleases, polymerases and ligase. In the case of multicellular eukaryotes such as humans the nuclease is recruited, Artemis:DNA PKCS (DNA-dependent protein kinase catalytic subunit). Artemis:DNA PKCS is able to endonucleolytically cut a range of dirty DNA damaged ends through its 5' exonuclease, 3' and 5' endonuclease and hairpin opening activities, the same role is fulfilled by the MRX complex in yeast. Once the damaged DNA ends have been resected, this provides 3' overhangs for the polymerases to initiate DNA synthesis (Lieber 2010). In mammals, Pol  $\lambda$  and  $\mu$  are responsible for DNA synthesis during NHEJ, and Pol4 in *S. cerevisiae*, all polymerases are template-independent (Daley, Palmboos *et al.* 2005). In order for the repaired DNA ends to be ligated they are held in physical proximity by Ku:DNA PKCS complex (Downs and Jackson 2004) and ligated by XLF:XRCC4:DNA ligase IV in mammals and Nej1:Lif1:Dnl4 in *S. cerevisiae*. XLF:XRCC4:DNA ligase IV is able to ligate across gaps, ligate incompatible ends and ssDNA, and can ligate one strand whilst the other has a complex configuration, for example bearing flaps, which can lead to a high number of errors (Lieber 2010). NHEJ plays a vital role in V(D)J

recombination in mammals, generating genetic diversity in immune B-cell and T-cell proteins (Jung and Alt 2004).

#### **1.4.17 Non-homologous end joining in *E. coli***

As in eukaryotes, a DSB is recognized by the protein Ku, which binds DNA to form Ku:DNA complex. Ku acts as a toolbelt for the multidomain protein LigD. LigD possesses a ligase domain LigD, a polymerase domain PolDom and a nuclease domain NucDom. LigD is recruited to a DSB by Ku via the PolDom domain, where a synapsis between the two DNA:Ku:LigD complexes is formed. The NucDom resects the 5' ends of the damaged DNA to generate 3' overhangs. PolDom then aligns the 3' ssDNA overhangs and fills the remaining gaps by 3'-end extension. PolDom is able to dislocate and realign the 3' overhangs, which can generate base substitutions and frame-shift deletions. PolDom possesses the capacity to recognize a variety of NHEJ DNA intermediates, but has a low DNA synthesis rate and does not possess a high fidelity. The repaired DNA ends are held in close proximity to allow ligation by LigD (Pitcher, Brissett *et al.* 2007).





**Figure 1.33 Non-homologous end joining in *E. coli*.**

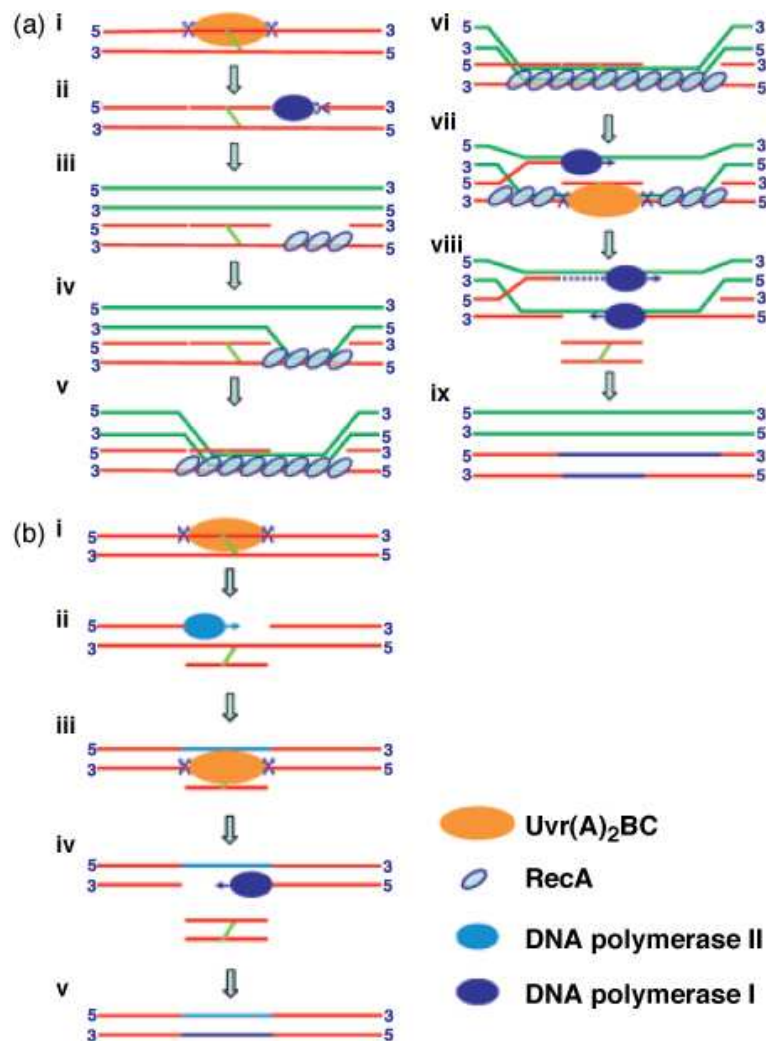
Recognition and binding by Ku leads to the recruitment of LigD followed by synapse formation. Nuclease resection by NucDom generates 3' overhangs that are then aligned by PolDom. Resynthesis by PolDom and finally ligation of nicks concludes NHEJ repair. Taken from (Pitcher, Brissett *et al.* 2007).

#### **1.4.18 Interstrand cross-link repair**

Interstrand cross-links (ICLs) are DNA lesions that covalently link both DNA strands, the failure to remove these ICLs can result in the blocking of DNA replication and transcription due to the inability to separate the DNA strands, and therefore ultimately cell death (McCabe, Olson *et al.* 2009) (Hlavin, Smeaton *et al.* 2010). A variety of bifunctional alkylating agents including mitomycin C react with DNA to create ICLs. ICLs can also arise from the by-products of metabolism, for example from the lipid peroxidation product malondialdehyde (McCabe, Olson *et al.* 2009).

#### **1.4.19 Interstrand cross-link repair in *E. coli***

ICL repair in *E. coli* is performed both by NER and homologous recombination (HR) proteins as shown in Figure 1.34. The NER enzymes UvrABC incise one strand both 5' and 3' of the ICL, the exonuclease activity of DNA polymerase I creates a ssDNA gap on the 3' side of the ICL, allowing the recombinase RecA to form a nucleofilament on the ssDNA, and pair with intact homologous dsDNA. Strand exchange past the ICL oligonucleotide is catalysed by RecA and the region of homologous pairing is extended. DNA synthesis to fill the gap is initiated by DNA polymerase I using the 3' end. The resulting Holliday junctions are resolved by the Holliday junction resolvase RuvABC. UvrABC incises on both sides of the ICL, the resulting gap is filled by DNA polymerase I and the nick is sealed by DNA ligase (Lehoczky, McHugh *et al.* 2007).



**Figure 1.34 Interstrand cross link repair pathway in *E. coli***

(a) Repair pathway of an ICL using NER and HR. i) ICL is recognized by UvrA and bound by UvrABC endonuclease, which then proceeds to incise on both the 5' and 3' sides of the ICL, generating a 13 nucleotide fragment that is covalently bound to the complementary DNA strand via the ICL. ii) The incised strand is resected by the 5'-exonuclease activity of DNA polymerase I (Pol I). iii) RecA polymerises into the resulting ssDNA and iv) catalyses strand exchange with homologous duplex DNA. v, vi) the region of strand exchange is extended to include the ICL. vii) Pol I extends the 3' overhang using the homologous donor DNA as a template. UvrABC again incises either side of the ICL. viii) The resulting ssDNA gap is filled by Pol I and remaining nicks sealed by DNA ligase. (b) HR independent ICL repair. i) ICL is recognized by UvrA and bound by UvrABC endonuclease, which then proceeds to incise on both the 5' and 3' sides of the ICL, generating a 13 nucleotide fragment unhooking the ICL. ii) The resulting gap is filled by DNA polymerase II. iii) UvrABC then incises again either side of the ICL, removing the ICL fragment from the duplex DNA. iv) The remaining gap is filled in by DNA polymerase II, and the remaining nicks sealed by DNA ligase. Taken from (Lehoczky, McHugh *et al.* 2007).

Alternatively repair of ICLs can be achieved by the NER and DNA polymerase II-dependent translesion pathways, shown in Figure 1.34. As before, UvrABC incises on

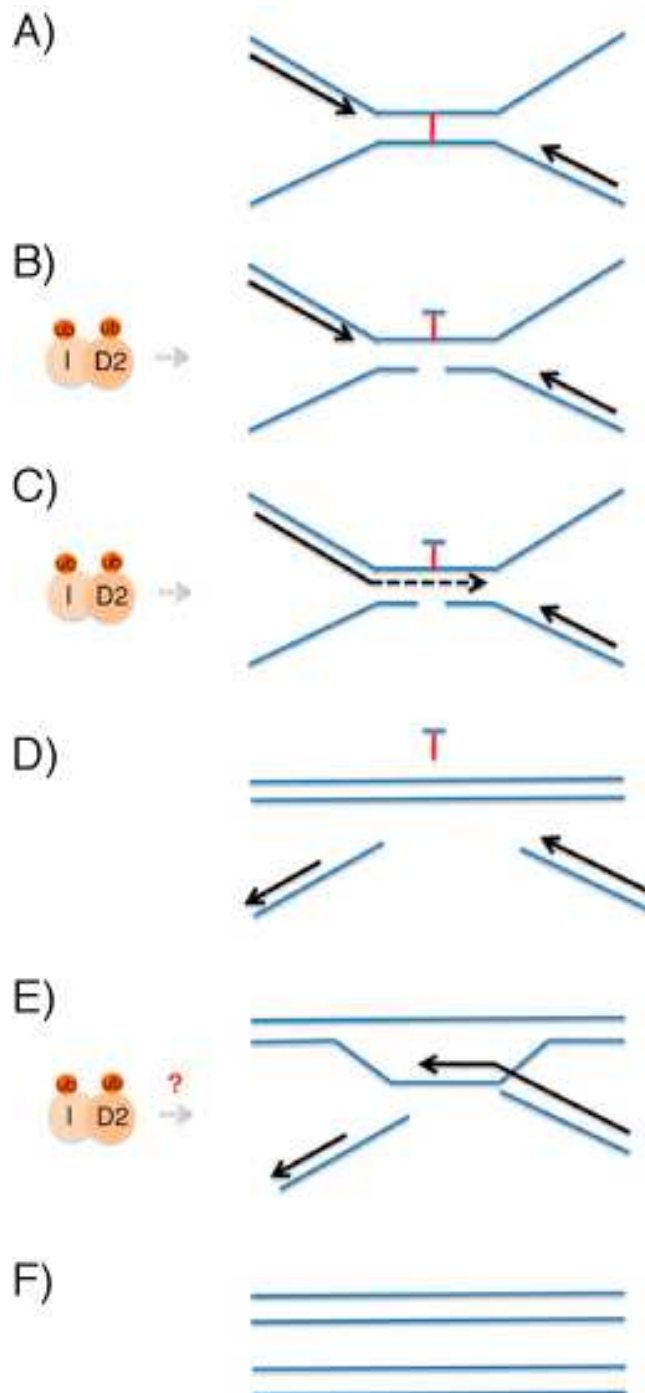
both side of the ICL. This is followed by translesion dependent synthesis by DNA polymerase II, thereby unhooking the ICL oligo shown in b(ii) of Figure 1.34. Again UvrABC incises both 5' and 3', removing the ICL containing DNA fragment. Translesion synthesis by polymerase II fills the resulting gap and the remaining nick is sealed by DNA ligase (Lehoczky, McHugh *et al.* 2007).

#### **1.4.20 Interstrand cross-link repair in eukaryotes**

In eukaryotes, ICL repair involves NER and HR pathways and the formation of double-stranded DNA breaks (DSB). The NER endonuclease XPF-ERCC1 is thought to recognize and incise ICLs in dsDNA, creating a DSB; the Mus81-Eme1 complex is also able to incise at ICLs. The resulting DSB intermediate is then repaired by the HR pathway, which is discussed elsewhere. However, if an ICL is recognized and incised but not removed during DNA replication, this can lead to a stalled replication fork. The NER pathway removes the ICL and the resulting DSB is repaired by HR, allowing the replication fork to be reloaded and DNA replication to continue (Lehoczky, McHugh *et al.* 2007; McCabe, Olson *et al.* 2009).

The Fanconi anemia (FA) pathway also co-operates with DNA repair proteins to resolve ICLs during DNA replication. A large nuclear E3 ubiquitin ligase complex, known as the FA core complex, consists of FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, FANCL and FANCM proteins, it monoubiquitinates the FANCD2/FANCI heterodimer. This results in the stimulation of repair of the DSB intermediate by HR, instead of the error prone non-homologous end joining (NHEJ) method, via recruitment of the newly identified FAN1 nuclease, (whose exact role is still unknown) and DNA repair proteins including BRAC2. The FANCM protein, homolog of the archaeal Hef, has also been shown to have a FA pathway-independent role in activating check point signalling and remodelling of replication forks via branch migration to assist ICL repair (Kee and D'Andrea 2010). Mutation of FA proteins manifests in FA, a multigenetic disorder that can be classified into thirteen complementation groups and is characterised by chromosome instability (Hlavin, Smeaton *et al.* 2010). FA patients are highly sensitive to ICL inducing agents due to defective ICL repair by HR. As a consequence, ICL remain during DNA replication and cause a chromatid break, which is repaired by HR between non-homologous DNA sequences on a non-sister chromatid. This illegitimate HR is able to occur due to the

lack of tight regulation by FA proteins and results in chromatid fusion known as a radial fusion. The treatment of FA patients with ICL inducing agents yields chromatid breaks and radial structures in 30% - 100% of the cells due to illegitimate HR (Niedernhofer, Lalai *et al.* 2005).

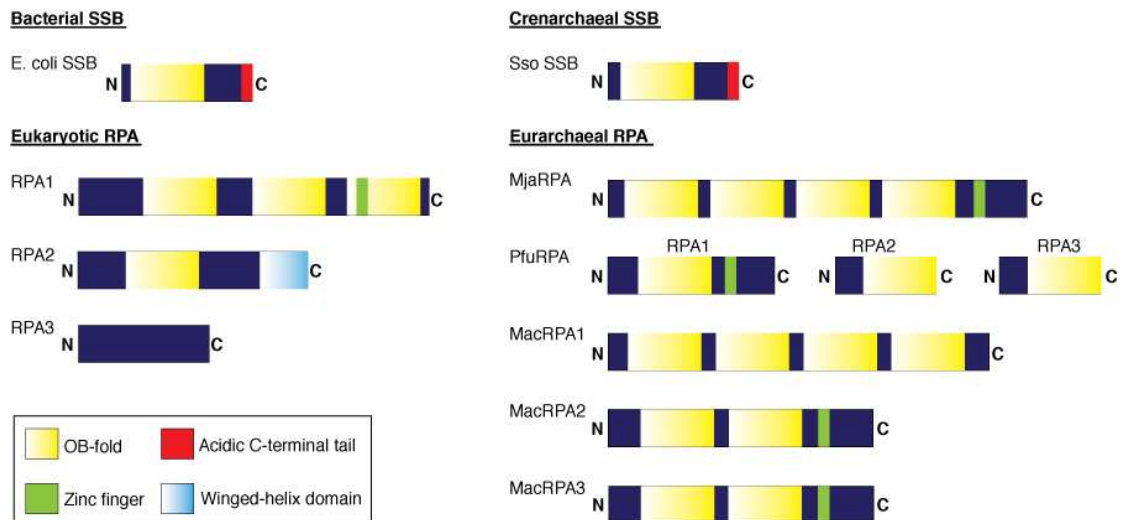


**Figure 1.35 ICL repair in eukaryotes**

A) Presence of an ICL stalls replication fork, triggering the monoubiquitination of the FANCD2/I heterodimer. B) ICL is incised by XPF-ERCC1 and Mus81-EME1, or the recently identified FAN1, unhooking the ICL. C) The resulting gap is bypassed by a translesion polymerase, for example REV1. D) The ICL is completely removed from the DNA by XPF-ERCC1 and Mus81-EME1 activity and the resulting gap filled in by DNA polymerase  $\delta$  or  $\epsilon$ . E) The resulting DSB is repaired by homologous recombination. F) Repaired duplex DNA Taken from (Kee 2010)

## 1.5 Single stranded DNA binding proteins

Single-stranded DNA binding proteins bind to ssDNA with high affinity, and to dsDNA and RNA with low affinity, (Wobbe, Weissbach *et al.* 1987) (Kim, Snyder *et al.* 1992) (Wold, Weinberg *et al.* 1989): they play a vital organizational role in the central genome maintenance of the cell, providing docking platforms for a wide range of enzymes to gain access to genomic substrates (Lu and Keck 2008). The bacteriophage T4 gene 32 monomer was the first single-stranded DNA binding protein to be identified (Alberts and Frey 1970). The single-stranded DNA binding protein is denominated SSB in bacteria and replication protein A (RPA) in eukaryotes. RPA was first identified as an essential protein for DNA replication in the eukaryotic simian virus (SV40) (Wobbe, Weissbach *et al.* 1987) by stimulating the T antigen-mediated unwinding of the SV40 origin of replication (Kenny, Lee *et al.* 1989). RPA and SSB have now been established as an essential protein for DNA metabolism including DNA replication, recombination and repair in all domains of life (Wobbe, Weissbach *et al.* 1987; Heyer, Rao *et al.* 1990; Coverley, Kenny *et al.* 1991; Moore, Erdile *et al.* 1991; Coverley, Kenny *et al.* 1992; Wold 1997). In order for any processes of DNA metabolism to occur the genomic DNA must be unwound leaving it vulnerable to nuclease and chemical attack, as well as open to the possibility of forming secondary structures. RPA and SSB binding to ssDNA prevents any of these events from occurring (Lu, Windsor *et al.* 2009). The basic architecture of RPA and SSB is based on the oligonucleotide/ oligosaccharide binding (OB) fold, a five-stranded  $\beta$ -sheet coiled into a closed barrel, but the number of OB-folds present varies from species to species (Bochkarev and Bochkareva 2004; Fanning, Klimovich *et al.* 2006).



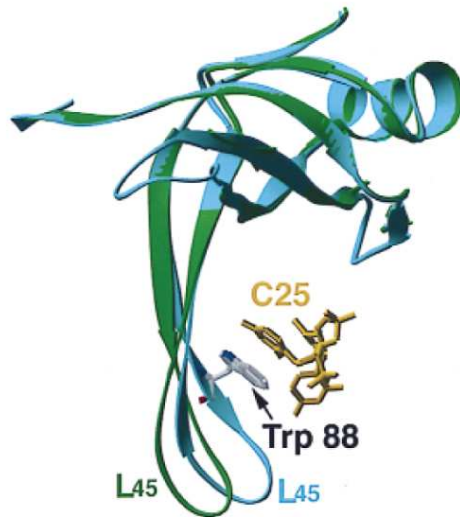
**Figure 1.36 The domain organization of RPAs and SSBs**

Domains structure of the single-stranded DNA binding proteins from eukaryotes (RPA), bacteria (*E. coli* SSB), crenarchaea (*Sulfolobus solfataricus*, Sso) and euryarchaea (*Methanocaldococcus jannaschii*, Mja), (*Pyrococcus furiosus*, Pfu) and (*Methanosarcina acetivorans*, Mac). Domains not drawn to scale.

### 1.5.1 Bacterial single-stranded DNA binding protein (SSB)

Bacterial SSB proteins use an OB-fold to bind ssDNA (Lohman and Ferrari 1994), the *Escherichia coli* SSB homotetramer (Sigal, Delius *et al.* 1972) is made up of 20 kDa subunits and has been shown to interact with proteins involved in DNA replication, repair and recombination (Shereda, Kozlov *et al.* 2008). The N-terminus of *E. coli* SSB contains a single OB-fold used to bind ssDNA and is responsible for tetramerisation (Meyer and Laine 1990) (Shereda, Reiter *et al.* 2009). The flexible C-terminus tail of *E. coli* SSB contains an amphipathic peptide sequence that terminates in a hydrophobic tripeptide and also possesses acidic properties, both of which are well conserved (Lu and Keck 2008). The C-terminus is required for protein-protein interactions, which *in vivo* has been shown to be essential for SSB function (Chase, L'Italien *et al.* 1984; Curth, Genschel *et al.* 1996). For example, *E. coli* SSB interacts with the replicative helicase DnaB that is involved in recruiting DNA polymerase III to the DNA replication fork (Shereda, Kozlov *et al.* 2008) (Shereda, Reiter *et al.* 2009). The C-terminus of *E. coli* SSB also interacts and stimulates the DNA helicase RecQ indirectly by binding to ssDNA resulting from RecQ unwinding activity, preventing additional ssDNA-RecQ complexes forming that may inhibit the helicase activity (Shereda, Reiter *et al.* 2009) (Shereda, Kozlov *et al.* 2008). To regulate

protein interactions with the acidic C-terminus of *E. coli* SSB, the C-terminus interacts with the positively charged N-terminus. On binding ssDNA a conformational change occurs, resulting in the C-terminus becoming more accessible for potential protein-protein interaction (Kozlov, Cox *et al.* 2010).

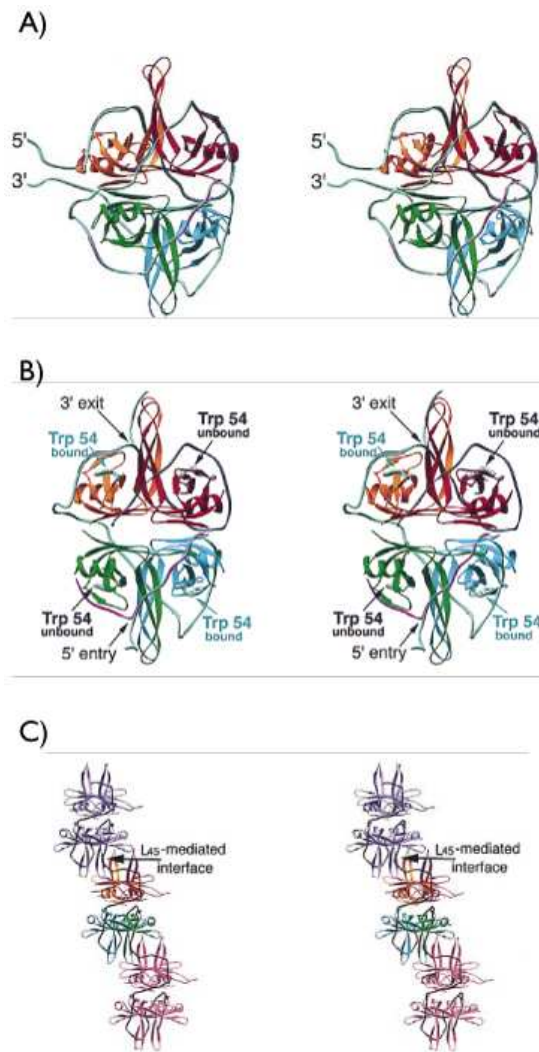


**Figure 1.37 *E. coli* SSB**

The structure of an *E. coli* SSB monomer bound to ssDNA (depicted in yellow) via the tryptophan amino acid labelled, in the same manner as in a tetramer. Cyan topped strands show the alpha-helix and the green looped strand show the beta-sheets. Taken from (Raghunathan, Kozlov *et al.* 2000).

Binding of *E. coli* SSB to ssDNA binding takes place in numerous modes depending on how many subunits are involved in ssDNA contacts, the two main modes are (SSB)<sub>35</sub> and (SSB)<sub>65</sub> where 35 and 65 nucleotides are occluded per bound SSB tetramer, respectively (Kozlov, Cox *et al.* 2010) (Lohman and Overman 1985). The transition between these ssDNA binding modes is reversible and is dependent on the salt concentration and SSB:DNA ratio, for example at high binding density the (SSB)<sub>35</sub> ssDNA binding mode is favoured (Lohman and Overman 1985) (Bujalowski, Overman *et al.* 1988).





**Figure 1.38 The ssDNA-binding modes of *E. coli* SSB.**

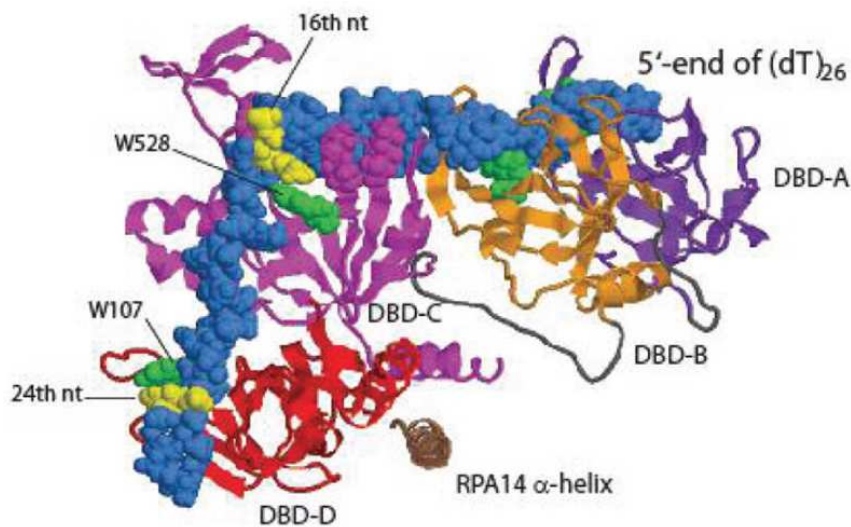
A) Model of the *E. coli* SSB<sub>65</sub> ssDNA-binding mode, each monomer is coloured individually; orange, magenta, cyan and green. The ssDNA is depicted blue-green and the 5' and 3' ends have been extended for clarity. B) Model of the *E. coli* SSB<sub>35</sub> ssDNA-binding mode. The point of ssDNA 5' entry and 3' exit, in addition to the tryptophan residues bound and unbound by ssDNA are indicated. C) Interaction between SSB tetramers via the L<sub>45</sub> loop, allowing assembly of SSB tetramers during the SSB<sub>35</sub> ssDNA-binding mode. Taken from (Ragunathan, Kozlov *et al.* 2000).

The Gram-positive bacterium *Streptococcus pneumoniae* has two ssDNA-binding proteins SsbA and SsbB, SsbA has a size of 17.3 kDa and SsbB has a slightly smaller size of 14.9 kDa (Grove, Willcox *et al.* 2005). Both SsbA and B form tetramers however SsbA and SsbB have very different DNA-binding activities; SsbA is able to bind ssDNA in both (SSB)<sub>35</sub> and (SSB)<sub>65</sub> modes whereas SsbB is not (Grove, Willcox *et al.* 2005). The finding that SsbB expression is induced during natural transformation suggests that SsbB has a specific function whereas SsbA is the main ssDNA-binding protein in the cell (Peterson, Sung *et al.* 2004).

### **1.5.2 Eukaryotic RPA structure and function**

The eukaryotic RPA is a heterotrimeric complex consisting of 70 kDa, 32 kDa and a 14 kDa subunits known as RPA1, RPA2 and RPA3, respectively (Wold 1997). Haploinsufficiency of RPA1 results in defective G2-M checkpoint regulation, which

is seen in human disorders including Seckel syndrome, Microcephalic primordial dwarfism type II and Nijmegen breakage syndrome. The 17p13.3 duplication syndrome arises through overexpression of RPA1 by gene duplication. Elevated levels of RPA1 have an adverse impact on HR repair of chromatid breaks, which result in chromosomal fusions (Outwin, Carpenter *et al.* 2011). The structure of RPA1 consists of a N-terminus, a central DNA binding domain containing two DNA binding domains DBDA and DBDB as well as a C-terminus containing a third DNA binding domain DBDC (Wold 1997). The N-terminus of RPA1 plays a major role in RPA-protein interactions required for all aspects of DNA metabolism, for example interaction with p53, DNA polymerase  $\alpha$  and the xeroderma pigmentosum damage-recognition protein (XPA) that acts in the DNA repair pathway nucleotide excision repair (NER) (Braun, Lao *et al.* 1997; Wold 1997). The central DNA binding domain is essential for RPA function, containing OB-folds DBDA and DBDB, which mediate the first stage of ssDNA binding (Bochkareva, Korolev *et al.* 2002). There is also evidence that suggests DBDA and DBDB, in addition to the N-terminus, are involved in RPA-protein interactions (Wold 1997) (Jacobs, Lipton *et al.* 1999). For example, direct interaction between the central binding domain and XPA have been shown (Walther, Gomes *et al.* 1999). In addition both the N-terminal domain and the central binding domain of RPA70 have been shown to be responsible for stimulating and increasing the processivity of DNA polymerase  $\alpha$  through both RPA-protein interactions and ssDNA-binding activity (Braun, Lao *et al.* 1997). The C-terminus of RPA1 contains a third DNA binding domain, DBDC, which contains a zinc-finger motif CX<sub>4</sub>CX<sub>14</sub>CX<sub>2</sub>C. The presence of this zinc finger is conserved in all known RPA1 homologues (Wold 1997; Bochkareva, Korolev *et al.* 2000). The C-terminus of RPA1 has been shown to interact with RPA2 and the zinc finger was originally thought to be functionally and structurally important to ensure the formation of a RPA heterotrimer (Wold 1997). Later mutational analysis of the zinc finger provided conflicting results as to whether the RPA was still active in DNA replication and base excision repair (BER) when the zinc finger had been mutated or deleted (Lin, Chen *et al.* 1996; Lin, Shivji *et al.* 1998). More recently, the zinc motif has been shown to interact directly with DNA, in addition mutation of the zinc finger results in a structural change of the RPA complex, a reduced ssDNA-binding activity of RPA and no RPA activity during DNA replication (Walther, Gomes *et al.* 1999)



**Figure 1.39 Eukaryotic RPA binding to single-stranded DNA**

A model for the heterotrimeric eukaryotic RPA bound to 26 nucleotides of single-stranded DNA (dT)<sub>26</sub>, coloured blue. The DNA-binding domains (DBD) A, B, C, and D are coloured purple, orange, magenta and red. The DBD-A-C are located in the RPA70 subunit and the DBD-D in RPA32, as shown in Figure 1.36. The linker regions joining DBD-A to DBD-B and DBD-B to DBD-C are shown in grey and the  $\alpha$ -helix of RPA14 is coloured brown. Four Tryptophan residues, coloured green, are thought to interact with ssDNA, specifically dT16 and dT24, which are highlighted in yellow. Taken from (Cai, Roginskaya *et al.* 2007).

RPA2 possesses one OB-fold, DNA binding domain D (DBD-D), in addition to a flexible N-terminus and a C-terminus involved in protein-protein interactions (Wold 1997). The C-terminus of RPA2 contains a winged helix-loop-helix domain that has been shown to bind DNA repair proteins once RPA2 has bound to site of lesion. For example, interaction between the C-terminus of RPA2 and XPA assists XPA interaction with damaged DNA site. Once a stable RPA-XPA complex has been formed on the damaged DNA, other nucleotide excision repair (NER) proteins such as the endonuclease XPG-ERCC1 can be recruited (He, Henricksen *et al.* 1995; Stigger, Drissi *et al.* 1998) (Matsunaga, Park *et al.* 1996). The central DNA binding domain of RPA2 is critical for ssDNA binding and has been shown to be essential for RPA to support NER, suggesting that a mutant RPA lacking ssDNA binding ability cannot form a complex with XPA on DNA (Stigger, Drissi *et al.* 1998). In addition, the C-terminus of RPA2 has been shown to interact with RAD52, a protein involved in homologous recombination, *in vitro* and *in vivo* (Park, Ludwig *et al.* 1996). Interaction between C-terminus of RPA2 and the uracil-DNA glycosylase (UNG), an

essential enzyme of base excision repair (BER), has also been observed. One of two of the RPA2 binding regions of UNG is believed to recruit UNG to a DNA repair complex associated with a replication complex, to allow identification of uracil before DNA replication (Nagelhus, Haug *et al.* 1997; Otterlei, Warbrick *et al.* 1999). Additionally RPA2 has been suggested to play a role in UNG localization to replication foci to facilitate BER (Otterlei, Warbrick *et al.* 1999).

The N-terminus of RPA2 is phosphorylated to varying degrees during the cell cycle and in response to DNA damage. Binding of RPA2 to ssDNA induces a conformational change in RPA2, resulting in a substrate for phosphorylation by multiple kinases including ATR, ATM and DNA-dependent protein kinase (DNA-PK), all members of the phosphatidylinositol 3-kinase-like kinase family (PIKKs). Cell cycle-dependent phosphorylation of RPA2 occurs during the transition from G<sub>1</sub> to S phase, also resulting in a conformational change of RPA2 (Din, Brill *et al.* 1990). RPA2 phosphorylation stimulates DNA replication during initiation, for example at the unwinding of the replication origin, and during elongation (Dutta and Stillman 1992). Consequently, RPA2 phosphorylation is thought to be one of many regulatory steps of DNA replication (Din, Brill *et al.* 1990) (Dutta and Stillman 1992). During mitosis RPA2 undergoes phosphorylation by cyclin B-Cdk 1 on S23 and S29, which stimulates further hyperphosphorylation of RPA2 in response to DNA damage during mitosis but not during G<sub>1</sub> phase cells (Anantha, Sokolova *et al.* 2008)

RPA3 is folded into one structural domain of which the C-terminus plays a significant structural role in the trimerisation of the three subunits RPA1, 2 and 3 (Wold 1997; Bochkareva, Korolev *et al.* 2002). Interactions between RPA3 and ssDNA have been detected (Wold 1997) though RPA3 does not possess an OB-fold, so a role in ssDNA binding similar to that of RPA1 and 2 is unlikely (Fanning, Klimovich *et al.* 2006).

#### **RPA2 homolog RPA4**

A mammalian specific homolog of RPA2, RPA4, which shares a 47% amino acid homology with RPA2 has been discovered (Keshav, Chen *et al.* 1995). RPA4 has one DNA binding domain, DBDG, which is the region of highest similarity to RPA2 (Haring, Humphreys *et al.* 2010). RPA4 has been shown to interact and form a

complex with RPA1 and 3 but not with RPA2 (Keshav, Chen *et al.* 1995). The RPA1, 3 and 4 complex however has distinctive properties, suggesting RPA4 may interact with different proteins and have a different cellular role to RPA2 (Keshav, Chen *et al.* 1995).

In contrast to RPA2, RPA4 does not undergo cell cycle or DNA-damage dependent phosphorylation (Keshav, Chen *et al.* 1995), (Haring, Humphreys *et al.* 2010). RPA4 is expressed highly and specifically in non-proliferating cells and has the ability to inhibit cells progressing from G<sub>1</sub> into S phase of the cell cycle in the presence of RPA2, signifying the possibility that RPA4 expression is related to cells entering quiescence (Keshav, Chen *et al.* 1995) (Haring, Humphreys *et al.* 2010).

### **1.5.3 Yeast RFA**

*Saccharomyces cerevisiae* possesses three single-stranded DNA binding proteins denominated Replication Factor A (RFA) 1, 2 and 3, which consist of a 69, 36 and 13 kD subunits respectively (Brill and Stillman 1991). All three individual RFAs have been shown to be essential for DNA replication, recombination and repair in *S. cerevisiae* (Longhese, Plevani *et al.* 1994). The three RFAs form a multi-subunit complex and during homologous recombination RFA1 binds and stabilizes ssDNA before displacement by the recombinase Rad51, as in higher eukaryotes (Brill and Stillman 1991) (Kantake, Sugiyama *et al.* 2003).

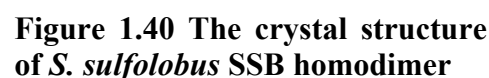
Phosphorylation of *S. cerevisiae* RFA2, as with hRPA2, occurs in a cell-cycle dependent manner on entry into S-phase of the cell-cycle and could be a regulatory method helping to ensure DNA replication only occurs in the S-phase of the cell cycle (Din, Brill *et al.* 1990). Investigations where the use of monoclonal antibodies against RFA2 have shown inhibition of DNA replication thus supporting the theory that despite only RFA1 possessing ssDNA binding activity, RFA2 and RFA3 may play an essential role in the cell (Erdile, Wold *et al.* 1990).

### **1.5.4 Archaeal RPA**

The domains of euryarchaeal and crenarchaeal ssDNA binding proteins are shown in Figure 1.36. Between crenarchaea and euryarchaea there is a wide diversity of single-stranded DNA binding proteins. Crenarchaea possess SSBs similar to those of bacteria,

consisting of a single subunit with one OB-fold and an acidic C-terminus tail (Rolfmeier and Haseltine 2010). Euryarchaea have RPA-like proteins that show homology with the eukaryotic RPA, but from species to species the architecture of euryarchaeal RPAs varies dramatically from a single polypeptide RPA to an RPA made up of several subunits. Each of these RPAs can contain up to four OB-folds as well as a zinc finger motif (Rolfmeier and Haseltine 2010).

*Sulfolobus solfataricus* has a bacterial-like SSB consisting of a small 20 kDA peptide containing one OB-fold and an acidic C-terminus tail (Haseltine and Kowalczykowski 2002; Rolfmeier and Haseltine 2010). The *S. solfataricus* SSB quaternary structure is similar to that of *E. coli* SSB, however the primary structure of the OB-fold shows greater homology to that of the eukaryotic RPA70 DBDB. This suggests that crenarchaeal SSBs may be structurally similar to bacterial SSB but at a nucleotide sequence level show homology to the eukaryotic RPA (Haseltine and Kowalczykowski 2002) (Kerr, Wadsworth *et al.* 2003). The acidic C-terminal tail of *S. solfataricus* provides a platform for protein-protein interactions, for example direct interaction between the SSB C-terminal tail and RNA polymerase has been identified independent of DNA. Furthermore, this interaction has been shown to stimulate transcription activity of the RNA polymerase (Richard, Bell *et al.* 2004). In *S. solfataricus* there is an absence of DNA damage recognition proteins such as homologs of XPA or XPC to initiate NER. The ability of *S. solfataricus* SSB to specifically bind and melt damaged duplex DNA *in vitro* suggests SSB may play a role in the identification and binding of damaged DNA, followed by the subsequent recruitment of NER repair proteins (Cubeddu and White 2005).



The ball-and-stick structures represent sulphate ions. The interaction between the L12 loops and the sulphate ions is thought to mimic the interaction with DNA. The black arrow indicates the OB-fold of the lower subunit. Taken from (Kerr, Wadsworth *et al.* 2003).

80

of MCM and potentially RPA for extreme conditions and specialized roles (Chia, Cann *et al.* 2010).

*Methanosarcina acetivorans* possess three RPA subunits, MacRPA1, 2 and 3, however it is unlikely that the MacRPAs form a heterotrimeric complex as seen in *P. furiosus* and in eukaryotes. MacRPA1 contains four DNA-binding domains (DBD) containing OB-folds, DBDA-D but does not possess a zinc finger. MacRPA2 and 3 both have two OB-fold containing DBDs and a zinc finger. However, despite showing a similar domain structure, MacRPA2 is unable to bind ssDNA as strongly as MacRPA3. All three MacRPAs are able to stimulate primer extension by *M. acetivorans* DNA polymerase BI. This demonstrates an element of redundancy between the three MacRPAs and that they are unlikely to function as a heterotrimer (Robbins, Murphy *et al.* 2004).



## Chapter 2: Materials and Methods

### 2.1 Materials

#### 2.1.1 Tables

**Table 2.1** *Haloferax volcanii* strains.

< > denotes episomal plasmid and [ ] plasmid integrated into the chromosome.

Strain	Parent	Genotype	Notes
H164	H148	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA	Made by Thorsten Allers.
H195	H181	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB	(Guy, Haldenby <i>et al.</i> 2006)
H678	(WT)		Wild-type isolate DS2 (Mullakhanbhai and Larsen 1975).
H888	H164	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1	Linear transformation with BstXI/BamHI trpA fragment of p933, now trpA+.
H891	H195	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB rpa2+:[ $\Delta$ rpa2 pyrE2+]	Integration of p934 ( $\Delta$ rpa2, pyrE2+) at rpa2 locus.
H912	H888	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 <p.tnaA::rpa2 pyrE2+>	Contains p938 expression vector for tryptophan-inducible over expression of rpa2.
H914	H888	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 <p.tnaA::6Xhis::rpa2 pyrE2+>	Contains p939 expression vector for tryptophan-inducible over expression of his-tagged rpa2.
H936	H195	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB rpe+:[ $\Delta$ rpe pyrE2+]	Integration of p944 ( $\Delta$ rpe, pyrE2+) at rpe locus.
H1025	H195	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB rpa2+:[ $\Delta$ rpa2::trpA+ pyrE2+]	Integration of p1003 ( $\Delta$ rpa2::trpA+, pyrE2+) at rpa2 locus.
H1026	H292	$\Delta$ pyrE2 bgaHa-Kp $\Delta$ trpA rpa2+:[ $\Delta$ rpa2::trpA+ pyrE2+]	Integration of p1003 ( $\Delta$ rpa2::trpA+, pyrE2+) at rpa2 locus.
H1028	H98	$\Delta$ pyrE2 $\Delta$ hdrB <rpa2+ pyrE2+ hdrB+>	Contains overexpression of rpa2 p1013 shuttle vector (rpa2+, pyrE2+, hdrB+)

H1038	H195	$\Delta pyrE2$ <i>bgaHa-Bb::[bgaHa pyrE2+]</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$	Integration of p564 ( <i>bgaHa</i> , <i>pyrE2+</i> ) at <i>bgaHa-Bb</i> locus.
H1040	H1038	$\Delta pyrE2$ <i>bgaHa</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$	Pop-out of p564 ( <i>bgaHa</i> , <i>pyrE2+</i> ) from <i>bgaHa-Bb</i> locus to generate <i>bgaHa</i> .
H1134	H936	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ $\Delta rpe$	Pop-out of p944 ( $\Delta rpe$ , <i>pyrE2+</i> ) at the <i>rpe</i> locus to generate $\Delta rpe$ .
H1190	H1134	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ $\Delta rpe$ <i>mre11rad50+::[<math>\Delta mre11rad50</math> pyrE2+]</i>	Integration of p199 ( $\Delta mre11$ <i>rad50</i> , <i>pyrE2+</i> ) at <i>mre11rad50</i> locus to generate triple $\Delta rpe$ $\Delta mre11$ <i>rad50</i> deletion.
H1191	H1134	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ $\Delta rpe$ <i>mre11+::[<math>\Delta mre11</math> pyrE2+]</i>	Integration of p141 ( $\Delta mre11$ , <i>pyrE2+</i> ) at <i>mre11</i> locus to generate double $\Delta rpe$ $\Delta mre11$ deletion.
H1193	H1025	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ <i>rpa2+::[<math>\Delta rpa2::trpA+</math> pyrE2+]</i> $\Delta rpa2+$ <i>hdrB+</i> <i>pyrE2+</i>	Episomal plasmid marked by <i>hdrB+</i> , <i>pyrE2+</i> with the WT <i>rpa2</i> gene under the control of its own promoter as a control for H1194 for complementation of <i>rpa2</i> .
H1198	H195	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ <i>rpap1+::[<math>\Delta rpap1::trpA+</math> pyrE2+]</i>	Integration of p1171 ( $\Delta rpap1::trpA+$ , <i>pyrE2+</i> ) at <i>rpap1</i> locus, to generate a $\Delta rpap1$ deletion.
H1200	H195	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ <i>rpa3+::[<math>\Delta rpa3::trpA+</math> pyrE2+]</i>	Integration of p1175 ( $\Delta rpa3::trpA+$ , <i>pyrE2+</i> ) at <i>rpa3</i> locus to generate a $\Delta rpa3$ deletion.
H1201	H195	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ <i>rpap1+::[<math>\Delta rpap1::trpA+</math> pyrE2+]</i>	Integration of p1176 ( $\Delta rpap1::trpA+$ , <i>pyrE2+</i> ) at <i>rpap1</i> locus to generate a $\Delta rpap1$ deletion.
H1209	H1205	$\Delta pyrE2$ $\Delta hdrB$ <i>Nph-pitA</i> $\Delta mrr$	Pop-out of p1160 ( $\Delta mrr$ , <i>pyrE2+</i> ) from <i>mrr</i> locus to leave $\Delta mrr$
H1214	H1190	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ $\Delta rpe$ $\Delta mre11rad50$	Pop-out of p199 ( $\Delta mre11$ , $\Delta rad50$ , <i>pyrE2+</i> ) from <i>mre11</i> locus, leaving behind $\Delta mre11$ and $\Delta rad50$ .
H1215	H1191	$\Delta pyrE2$ <i>bgaHa-Bb</i> <i>leuB-Ag1</i> $\Delta trpA$ $\Delta hdrB$ $\Delta rpe$ $\Delta mre11$	Pop-out of p141 ( $\Delta mre11$ , <i>pyrE2+</i> ) from <i>mre11</i> locus, leaving behind $\Delta mre11$ .

H1216	H1198	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpap1::trpA+	Pop-out of p1171 ( $\Delta$ rpap, pyrE2+, trpA+) from the rpa1ap locus to generate $\Delta$ rpap.
H1217	H1201	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpap::trpA+	Pop-out of p1176 ( $\Delta$ rpa1, pyrE2+, trpA+) at the rpa1 locus to generate $\Delta$ rpa1.
H1234	H729	$\Delta$ hdrB pyrE2+::[ $\Delta$ pyrE2::p.tna-rpa2 NovR]	Integration of p1187 (pyrE2::rpa2) at ectopic pyrE2 locus to insert rpa2 under an inducible tryptophanase promoter.
H1244	H1200	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpa3::trpA+	Pop-out of p1175 ( $\Delta$ rpa3::trpA+, pyrE2+) at rpa3 locus, leaving behind $\Delta$ rpa3::trpA+.
H1245	H195	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB rpa1 operon+::[ $\Delta$ rpa1 operon pyrE2+]	Integration of p1191 ( $\Delta$ rpa1 operon pyrE2+) at rpa1 operon locus.
H1246	H1245	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpa1 operon	Pop-out of p1191, generating $\Delta$ rpa1.
H1247	H1234	$\Delta$ hdrB $\Delta$ pyrE2::p.tna-rpa2	Pop-out of p1187 (pyrE2::rpa2), replacing pyrE2 with rpa2, to try and delete native rpa2.
H1257	H195	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB rpa3 operon+::[ $\Delta$ rpa3 operon::trpA+, pyrE2+]	Pop-in of p1210 ( $\Delta$ rpa3 operon, trpA+ pyrE2+) at rpa3 operon locus.
H1260	H1257	$\Delta$ pyrE2 $\Delta$ hdrB bgaHa-Bb $\Delta$ rpa3 operon::trpA+ leuB-Ag1 $\Delta$ trpA	Pop-out of p1210 ( $\Delta$ rpa3 operon trpA+ pyrE2+) at rpa3 operon locus, leaving behind $\Delta$ rpa3 trpA+.
H1261	H1209	$\Delta$ pyrE2 $\Delta$ hdrB Nph-pitA $\Delta$ mrr <p.tnaA::his tag-rpa1 pyrE2+ hdrB+>	Contains p1222 plasmid (pyrE2+ and hdrB+ markers) for expression of 6xHis tagged-rpa1 under the control of the tryptophan-inducible promoter.
H1262	H1209	$\Delta$ pyrE2 $\Delta$ hdrB Nph-pitA $\Delta$ mrr <p.tnaA::his tag-rpa1ap pyrE2+ hdrB+>	Contains p1223 plasmid (pyrE2+ and hdrB+ markers) for expression of 6xHis tagged-rpa1ap under a tryptophan-inducible promoter

H1280	H1286	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpap1	Pop-out of p1219 ( $\Delta$ rpap1ap, pyrE2+) from the $\Delta$ rpap1ap::trpA+ locus to generate $\Delta$ rpap1ap.
H1281	H1287	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpap1	Pop-out of p1197 ( $\Delta$ rpap1, pyrE2+) from the $\Delta$ rpap1::trpA+ locus to generate $\Delta$ rpap1.
H1282	H1246	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpap1 operon rpa3 operon+::[ $\Delta$ rpa3 operon::trpA+, pyrE2+]	Pop-in of p1210 ( $\Delta$ rpa3, pyrE2+, trpA+) at the rpa3 locus
H1283	H1209	$\Delta$ pyrE2 $\Delta$ hdrB Nph-pitA $\Delta$ mrr <p.tnaA::his tag-rpa3ap pyrE2+ hdrB+>	Contains N-terminally his- tagged rpa3ap under the tryptophan-inducible promoter
H1284	H1209	$\Delta$ pyrE2 $\Delta$ hdrB Nph-pitA $\Delta$ mrr <p.tnaA::6Xhis::rpa2 pyrE2+>	Contains N-terminally his- tagged rpa2 under tryptophan-inducible promoter.
H1285	H1209	$\Delta$ pyrE2 $\Delta$ hdrB Nph-pitA $\Delta$ mrr <p.tnaA::his tag-rpa3 pyrE2+ hdrB+>	Contains N-terminally his- tagged rpa3 under tryptophan-inducible promoter
H1286	H1216	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpap1::trpA+::[ $\Delta$ rpap1 pyrE2+]	Pop-in of p1219 ( $\Delta$ rpap1ap, pyrE2+) at the $\Delta$ rpap1ap::trpA+ locus.
H1287	H1217	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpap1::trpA+::[ $\Delta$ rpap1 pyrE2+]	Pop-in of p1197 ( $\Delta$ rpap1, pyrE2+) at the $\Delta$ rpap1::trpA+ locus.
H1326	H1280	$\Delta$ pyrE2 bgaHa-Bb leuB-Ag1 $\Delta$ trpA $\Delta$ hdrB $\Delta$ rpap1 rpa3+::[ $\Delta$ rpa3::trpA+ pyrE2+]	Pop-in of p1175 ( $\Delta$ rpa3::trpA+, pyrE2+) at rpa3 locus.
H1369	H1209	$\Delta$ pyrE2 $\Delta$ hdrB Nph-pitA $\Delta$ mrr <p.tnaA::his tag-rpa3 rpa3ap pyrE2+ hdrB+>	Plasmid p1280 (rpa3 operon with N-terminally his- tagged rpa3) was constructed in this strain, as toxic in <i>E. coli</i> .
H1370	H1209	$\Delta$ pyrE2 $\Delta$ hdrB Nph-pitA $\Delta$ mrr <p.tnaA::rpa3 his tag- rpa3ap pyrE2+ hdrB+>	Plasmid p1281 (rpa3 and N- terminally his-tagged rpa3ap) was constructed in this strain, as toxic in <i>E.</i> <i>coli</i> .

H1390	H1282	<i>ΔpyrE2 bgaHa-Bb leuB-Ag1 ΔtrpA ΔhdrB Δrpa1 operon rpa3 operon+::[Δrpa3 operon::trpA+, pyrE2+] &lt;rpa1 operon+hdrB+pyrE2&gt;</i>	Contains episomal complementary <i>rpa1</i> operon plasmid to generate double <i>Δrpa1 Δrpa3</i> operon deletion.
H1409	H195	<i>ΔpyrE2 bgaHa-Bb leuB-Ag1 ΔtrpA ΔhdrB rpa3ap+::[Δrpa3ap::trpA+ pyrE2+]</i>	Integration of p1285 ( <i>Δrpa3ap::trpA+ pyrE2+</i> ) at the <i>rpa3ap</i> locus
H1410	H1409	<i>ΔpyrE2 bgaHa-Bb leuB-Ag1 ΔtrpA ΔhdrB Δrpa3</i>	Pop-out of p1285 generating <i>Δrpa3ap</i>
H1424	H1422	<i>ΔpyrE2 ΔhdrB Nph-pitA Δmrr cdc48d-Ct</i>	Pop-out of p1294, replacing Hsal <i>cdc48d</i> with Hvo <i>cdc48d-Ct</i> (C-terminal truncation. Made by Thorsten Allers.
H1430	H1424	<i>ΔpyrE2 ΔhdrB Nph-pitA Δmrr cdc48d-Ct &lt;p.tna::his tag + pyrE2+ hdrB+&gt;</i>	Contains p963 plasmid ( <i>pyrE2+</i> and <i>hdrB+</i> markers), empty vector control for 6xHis tagged-protein expression. Made by Thorsten Allers.
H1470	H1209	<i>ΔpyrE2 ΔhdrB Nph-pitA Δmrr &lt;p.tnaA::his tag-rpa1 rpa1ap pyrE2+ hdrB+&gt;</i>	Contains p1326, his-tagged <i>rpa1</i> with <i>rpa1ap</i> downstream, was constructed in this strain, as toxic in <i>E. coli</i> .
H1471	H1209	<i>ΔpyrE2 ΔhdrB Nph-pitA Δmrr &lt;p.tnaA::rpa1 his tag-rpa1ap pyrE2+ hdrB+&gt;</i>	Contains p1327, his-tagged <i>rpa1ap</i> with <i>rpa1</i> inserted upstream, was constructed in this strain, as toxic in <i>E. coli</i> .
H1472	H1390	<i>ΔpyrE2 bgaHa-Bb leuB-Ag1 ΔtrpA ΔhdrB Δrpa1 operon Δrpa3 operon::trpA+</i>	Pop-out of p1210, generating double <i>Δrpa1</i> operon and <i>Δrpa3</i> operon deletion, and loss of complementary <i>rpa1</i> operon plasmid.
H1473	H1281	<i>ΔpyrE2 bgaHa-Bb leuB-Ag1 ΔtrpA ΔhdrB Δrpa1 rpa3ap+::[Δrpa3ap::trpA+ pyrE2+]</i>	Pop-in of p1285 at <i>rpa3ap</i> locus.
H1474	H729	<i>ΔhdrB pyrE2+::[hdrB pyrE2+]</i>	Integration of <i>hdrB</i> maker (p1293) at <i>pyrE2</i> locus, to check the replacement of <i>pyrE2</i> with another gene.

H1475	H1247	$\Delta hdrB \Delta pyrE2::p.tna-rpa2$ $rpa2+::[\Delta rpa2::p.fdx hdrB+$ $pyrE2+]$	Integration of the <i>rpa2</i> deletion construct marked <i>hdrB+</i> p1227.
H1476	H1247	$\Delta hdrB \Delta pyrE2::p.tna-rpa2$ $rpa2+::[\Delta rpa2::p.fdx hdrB+$ $pyrE2+]$	Integration of the <i>rpa2</i> deletion construct marked <i>hdrB+</i> p1227.
H1478	H1365	$\Delta pyrE2 \Delta hdrB bgaHa-Bb$ $\Delta rpa3$ operon <i>leuB-Agl</i> $\Delta trpA rpa1$ operon+:: $[\Delta rpa1$ operon:: <i>trpA+</i> $pyrE2+]$	Integration of <i>rpa1</i> operon deletion construct p1279, marked with <i>trpA</i> and episomal complementary <i>rpa1</i> operon plasmid p1272.
H1480	H1424	$\Delta pyrE2 \Delta hdrB Nph-pitA$ $\Delta mrr cdc48d-Ct$ < <i>p.tnaA::his tag-rpa3 rpa3ap</i> $pyrE2+ hdrB+$ >	p1280 with <i>rpa3ap</i> insert after N-terminally his-tagged <i>rpa3</i> . For over expression to examine <i>rpa3:rpa3ap</i> interaction.
H1481	H1424	$\Delta pyrE2 \Delta hdrB Nph-pitA$ $\Delta mrr cdc48d-Ct$ < <i>p.tnaA::rpa3 his tag-rpa3ap</i> $pyrE2+ hdrB+$ >	p1281 with <i>rpa3</i> inserted upstream of N-terminally his-tagged <i>rpa3ap</i> . For over expression to examine <i>rpa3:rpa3ap</i> interaction.
H1482	H1424	$\Delta pyrE2 \Delta hdrB Nph-pitA$ $\Delta mrr cdc48d-Ct$ < <i>p.tnaA::his tag-rpa1 rpa1ap</i> $pyrE2+ hdrB+$ >	p1326 containing his-tagged <i>rpa1</i> with <i>rpa1ap</i> inserted downstream. For over expression to examine <i>rpa1:rpa1ap</i> interaction.
H1483	H1424	$\Delta pyrE2 \Delta hdrB Nph-pitA$ $\Delta mrr cdc48d-Ct$ < <i>p.tnaA::rpa1 his tag-rpa1ap</i> $pyrE2+ hdrB+$ >	p1327 contains the <i>rpa1</i> inserted upstream of his-tagged <i>rpa1ap</i> . For over expression to examine <i>rpa1:rpa1ap</i> interaction.
H1491	H1424	$\Delta pyrE2 \Delta hdrB Nph-pitA$ $\Delta mrr cdc48d-Ct$ < <i>p.tnaA::his tag-rpa2 pyrE2+</i> $hdrB+$ >	Contains overexpression plasmid pTA981 containing N-terminally his-tagged <i>rpa2</i> .
H1493		$\Delta hdrB pyrE2+::[hdrB$ $pyrE2+]$	Pop-in of pTA1193 ( <i>pyrE2</i> deletion construct marked with <i>hdrB</i> )
H1494		$\Delta hdrB \Delta pyrE2::hdrB$	Pop-out of pTA1193, <i>pyrE2</i> deletion construct marked with <i>hdrB</i> , replacing <i>pyrE2</i> with <i>hdrB</i> .

**Table 2.2 *Escherichia coli* strains**

Strain	Genotype	Relevant properties
XL1-Blue MRF'	<i>endA1 gyrA96 (NalR) lac</i> [F' <i>proAB lacIqZΔM15 Tn10</i> ( <i>TetR</i> )] Δ( <i>mcrA</i> )183 Δ( <i>mcrCB-</i> <i>hsdSMR-mrr</i> )173 <i>recA1 relA1</i> <i>supE44 thi-1</i>	Standard cloning vector for blue/white selection using pBluescript based plasmids. Tetracycline used to select against loss of F' episome (Bullock, Fernandez <i>et al.</i> 1987).
N2338 (GM121)	<i>F- ara-14 dam-3 dcm-6</i> <i>fhuA31 galK2 galT22 hsdR3</i> <i>lacY1 leu-6 thi-1 thr-1 tsx-78</i>	<i>dam<sup>-</sup> dcm<sup>-</sup></i> mutant for preparing DNA for <i>H.</i> <i>volcanii</i> transformations (Allers, Ngo <i>et al.</i> 2004).

**Table 2.3 Plasmids**

Plasmid	Description
pBluescript II SK+	Standard <i>E. coli</i> cloning vector, with blue/white selection.
pGB68	1.7 kb flanking sequences of <i>H. volcanii</i> <i>pyrE2</i> for insertion of genes at the <i>pyrE2</i> locus, cloned into pBR-Nov cut with <i>Asp718</i> + <i>HindIII</i> (Bitan-Banin, Ortenberg <i>et al.</i> 2003)
pTA131	pBluescript II SK+ with 0.7 kb <i>BamHI/XbaI</i> <i>p.fdx::pyrE2</i> fragment inserted at <i>PsiI</i> site. (Allers, Ngo <i>et al.</i> 2004).
pTA298	Contains <i>trpA</i> marker between <i>NcoI</i> and <i>SphI</i> sites under a ferredoxin promoter from <i>H. salinarum</i> .
pTA356	Shuttle vector containing ampicillin and <i>hdrB</i> markers and ferredoxin promoter with 948bp <i>BmgBI-EcoRV</i> fragment from pTA250 (containing Hv oripHV1/4). Inserted at <i>PciI</i> site (Delmas, Shunburne <i>et al.</i> 2009).
pTA409	Shuttle vector containing ampicillin, <i>pyrE2</i> and <i>hdrB</i> markers and ferredoxin promoter with 948bp <i>BmgBI-EcoRV</i> fragment from pTA250 (containing Hv oripHV1/4) inserted at Klenow-blunted <i>PciI</i> site (Delmas, Shunburne <i>et al.</i> 2009).

---

pTA884	pBluescript with insertion of 5038 bp <i>EcoRI</i> to <i>NotI</i> genomic DNA fragment containing <i>rpa3</i> operon, inserted at <i>EcoRI</i> and <i>NotI</i> sites
pTA898	pBluescript with insertion of 7335 bp <i>EcoRI</i> to <i>NotI</i> genomic DNA fragment containing <i>rpa2</i> gene, inserted at <i>EcoRI</i> and <i>NotI</i> sites
pTA908	Shuttle vector containing ampicillin and <i>pyrE2</i> markers with 224 bp fragment of p.tnaA promoter, amplified from p870 and cut with <i>ApaI</i> and <i>ClaI</i> , inserted at <i>ApaI</i> and <i>ClaI</i> sites, for gene overexpression in <i>H. volcanii</i> .
pTA927	Overexpression vector with <i>pyrE2</i> marker and pHV2 origin, derived from pTA230 by insertion of 131-bp t.L11e terminator at <i>KpnI</i> site, 224-bp p.tnaA promoter at <i>ApaI</i> and <i>ClaI</i> sites, and 35-bp t.Syn terminator at <i>NotI</i> and <i>BstXI</i> sites (Allers, Barak <i>et al.</i> 2010).
pTA929	Overexpression vector with 6xHis tag, <i>pyrE2</i> marker, and pHV2 origin, derived from pTA927 by insertion of 26-bp fragment containing a His tag (CAC) <sub>6</sub> tract at the <i>NdeI</i> and <i>ClaI</i> sites (Allers, Barak <i>et al.</i> 2010).
pTA930	p927 expression shuttle vector containing <i>rpa2</i> coding sequence at restriction sites <i>NdeI</i> and <i>BamHI</i>
pTA931	p131 with <i>rpa2</i> deletion construct made by PCR with external primers <i>KpnI</i> and <i>XbaI</i> sites, internal primers with <i>BamHI</i> sites, inserted at <i>KpnI</i> and <i>XbaI</i> sites.
pTA934	p131 with <i>rpa2</i> deletion construct made by PCR with external primers <i>KpnI</i> and <i>XbaI</i> sites, internal primers with <i>BamHI</i> sites, inserted at <i>KpnI</i> and <i>XbaI</i> sites.
pTA936	p929 expression shuttle vector containing <i>rpa2</i> coding sequence at restriction sites <i>PciI</i> and <i>BamHI</i> .
pTA937	pBluescript with insertion of 8565 bp <i>BspEI</i> genomic DNA fragment containing <i>rpa1</i> operon, inserted at <i>XmaI</i> site.
pTA938	p927 expression shuttle vector containing <i>rpa2</i> coding sequence at restriction sites <i>NdeI</i> and <i>BamHI</i>

---



---

pTA939	p929 expression shuttle vector containing <i>rpa2</i> coding sequence at restriction sites <i>PciI</i> and <i>BamHI</i>
pTA940	p131 with <i>rpaI</i> -associated phosphoesterase ( <i>rpe</i> ) deletion construct made by PCR with external primers <i>KpnI</i> and <i>EcoRI</i> sites, internal primers with <i>BamHI</i> sites, inserted at <i>KpnI</i> and <i>EcoRI</i> sites.
pTA944	p131 with <i>rpaI</i> -associated phosphoesterase ( <i>rpe</i> ) deletion construct made by PCR with external primers <i>KpnI</i> and <i>EcoRI</i> sites, internal primers with <i>BamHI</i> sites, inserted at <i>KpnI</i> and <i>EcoRI</i> sites.
pTA955	p931 with <i>trpA BamHI</i> fragment from p298 inserted into <i>rpa2</i> deletion construct at <i>BamHI</i> site.
pTA962	Overexpression vector with <i>pyrE2</i> and <i>hdrB</i> markers and pHV2 origin, as well as a t.L11e terminator, p.tnaA promoter and t.Syn terminator(Allers, Barak <i>et al.</i> 2010).
pTA963	Overexpression vector with 6xHis tag, <i>pyrE2</i> and <i>hdrB</i> markers, and pHV2 origin, derived from pTA962 by insertion of a His tag (CAC) <sub>6</sub> tract(Allers, Barak <i>et al.</i> 2010).
pTA977	p963 containing N-terminal His tagged <i>rpa2</i> taken from p936 to investigate Rpa2 protein:protein interactions.
pTA981	p963 containing N-terminal His tagged <i>rpa2</i> taken from p936 to investigate Rpa2 protein:protein interactions.
pTA1003	p131 with <i>trpA BamHI</i> fragment from p298 inserted into <i>rpa2</i> deletion construct at <i>BamHI</i> site.
pTA1141	p131 containing <i>rpaI</i> deletion construct inserted at restriction sites <i>KpnI</i> and <i>XbaI</i> containing an internal <i>NdeI</i> site.
pTA1142	p131 containing <i>rpa3</i> deletion construct inserted at restriction sites <i>EcoRI</i> and <i>KpnI</i> containing an internal <i>NdeI</i> site.
pTA1165	pTA409 containing <i>rpa2</i> genomic clone for complementation when attempting to delete <i>rpa2</i> .

---

---

pTA1166	<i>rpa1</i> deletion construct pTA1141 with <i>trpA</i> marker, PCR amplified from pTA298 introducing <i>NdeI</i> restriction sites to insert at internal <i>NdeI</i> restriction site in pTA1141. Allowing selection for deletion over reversion to wild-type during pop-out process.
pTA1167	pTA409 containing <i>rpa2</i> genomic clone for complementation when attempting to delete <i>rpa2</i> .
pTA1168	pTA356 shuttle vector containing <i>rpa2</i> for complementation when attempting to delete <i>rpa2</i> , with additional <i>pyrE2</i> marker to ensure loss during pop-out process.
pTA1169	pTA356 shuttle vector containing <i>rpa2</i> for complementation when attempting to delete <i>rpa2</i> , with additional <i>pyrE2</i> marker to ensure loss during pop-out process.
pTA1170	Deletion construct of <i>rpa1ap</i> containing <i>trpA</i> marker from pTA298 inserted at <i>EcoRI</i> and <i>KpnI</i> sites in p131.
pTA1171	Deletion construct of <i>rpa1ap</i> containing <i>trpA</i> marker from pTA298 inserted at <i>EcoRI</i> and <i>KpnI</i> sites in p131.
pTA1174	<i>rpa3</i> deletion construct containing <i>trpA</i> marker from pTA1166 inserted at <i>NdeI</i> restriction site.
pTA1175	<i>rpa3</i> deletion construct containing <i>trpA</i> marker from pTA1166 inserted at <i>NdeI</i> restriction site.
pTA1176	<i>rpa1</i> deletion construct pTA1141 with <i>trpA</i> marker, PCR amplified from pTA298 introducing <i>NdeI</i> restriction sites to insert at internal <i>NdeI</i> restriction site in pTA1141. Allowing selection for deletion over reversion to wild-type during pop-out process.
pTA1184	pGB68 with <i>rpa2</i> under the tryptophanase promoter from pTA930, inserted using blunt ended ligation to effectively replace <i>pyrE2</i> with <i>rpa2</i> , to complement when trying to delete native <i>rpa2</i> .
pTA1185	<i>hdrB</i> ( <i>BspHI</i> - <i>SphI</i> PCR fragment) inserted at <i>NcoI</i> and <i>SphI</i> sites, under ferredoxin (fdx) promoter of <i>H. salinarum</i> . For making deletions in $\Delta$ <i>hdrB</i> background, <i>hdrB</i> marker is flanked by two <i>BamHI</i> sites. Made by Thorsten Allers.

---

---

pTA1187	pGB68 with <i>rpa2</i> under the tryptophanase promoter, inserted using blunt ended ligation to effectively replace <i>pyrE2</i> with <i>rpa2</i> , to complement when trying to delete native <i>rpa2</i> .
pTA1189	pTA131 with <i>rpa1</i> operon deletion construct inserted at restriction sites <i>XbaI</i> and <i>EcoRI</i> with an internal <i>NdeI</i> site.
pTA1191	pTA131 with <i>rpa1</i> operon deletion construct inserted at restriction sites <i>XbaI</i> and <i>EcoRI</i> with an internal <i>NdeI</i> site.
pTA1194	pTA931 <i>rpa2</i> deletion construct with <i>p.fdx-hdrB BamHI</i> fragment from pTA1185 inserted at <i>BamHI</i> site.
pTA1196	<i>rpa3</i> operon deletion construct, using <i>NdeI/EcoRI</i> downstream fragment from pTA1282 ( <i>rpa3ap</i> deletion construct) inserted at <i>NdeI/EcoRI</i> sites in pTA1142 ( <i>rpa3</i> deletion construct), to replace the downstream fragment of the <i>rpa3</i> deletion construct.
pTA1197	pTA131 containing <i>rpa1</i> deletion construct inserted at restriction sites <i>KpnI</i> and <i>XbaI</i> containing an internal <i>NdeI</i> site.
pTA1198	pTA131 containing <i>rpa3</i> deletion construct inserted at restriction sites <i>EcoRI</i> and <i>KpnI</i> containing an internal <i>NdeI</i> site.
pTA1206	Deletion construct of <i>rpa1</i> operon pTA1189 with <i>trpA</i> marker from pTA1166 inserted at <i>NdeI</i> site.
pTA1207	Deletion construct of <i>rpa3</i> operon pTA1196 with insertion of the <i>trpA</i> marker from pTA1166 at internal <i>NdeI</i> site.
pTA1210	Deletion construct of <i>rpa3</i> operon pTA1196 with insertion of the <i>trpA</i> marker from pTA1166 at internal <i>NdeI</i> site.
pTA1217	<i>rpa1ap</i> deletion construct pTA1170 with upstream and <i>trpA</i> fragment replaced with the upstream fragment amplified from pTA937 by PCR, to introduce compatible <i>SphI</i> sites, generating non- <i>trpA</i> marked deletion construct.
pTA1218	pTA963 with <i>rpa3</i> inserted downstream of His-tag. <i>AseI</i> inserted after <i>rpa3</i> stop codon to allow insertion of his-tagged <i>rpa3</i> upstream of his-tagged <i>rpa3ap</i> ( <i>AseI</i> is <i>NdeI</i> compatible).

---

---

pTA1219	<i>rpaIap</i> deletion construct pTA1170 with upstream and <i>trpA</i> fragment replaced with the upstream fragment amplified from pTA937 by PCR, to introduce compatible <i>SphI</i> sites, generating non- <i>trpA</i> marked deletion construct.
pTA1222	pTA963 with <i>rpaI</i> N-terminally his-tagged, has an <i>AseI</i> site downstream of <i>rpaI</i> to allow insertion of <i>rpaIap</i> as <i>NdeI</i> compatible.
pTA1223	pTA963 overexpression vector with <i>rpaIap</i> N-terminally His-tagged inserted at <i>PsiI</i> and <i>BamHI</i> sites. <i>rpaIap</i> was amplified by PCR from p937 introducing <i>BspHI</i> and <i>BamHI</i> sites.
pTA1224	PTA963 with <i>rpa3ap</i> N-terminally His-tagged inserted at <i>PsiI</i> and <i>EcoRI</i> sites. <i>Rpa3ap</i> was amplified by PCR from p884 introducing <i>BspHI</i> and <i>EcoRI</i> sites.
pTA1227	p931 <i>rpa2</i> deletion construct with <i>p.fdx-hdrB BamHI</i> fragment from p1185 inserted at <i>BamHI</i> site.
pTA1265	pTA409 overexpression plasmid with insertion of <i>rpaI</i> operon for complementation of deletions.
pTA1272	pTA409 overexpression plasmid with insertion of <i>rpaI</i> operon for complementation of deletions.
pTA1279	Deletion construct of <i>rpaI</i> operon pTA1189 with <i>trpA</i> marker from pTA1166 inserted at <i>NdeI</i> site, generating <i>trpA</i> marked <i>rpaI</i> operon deletion construct.
pTA1280	pTA1218 with <i>rpa3ap</i> amplified from pTA884 by PCR and inserted at <i>BstEII</i> and <i>EcoRI</i> sites after the N-terminally his-tagged <i>rpa3</i> , maintaining the reading frame.
pTA1281	pTA1224 with <i>rpa3</i> amplified from pTA884 by PCR and inserted upstream of N-terminally his-tagged <i>rpa3ap</i> at <i>NdeI</i> site.
pTA1282	<i>rpa3ap</i> deletion construct with upstream and downstream regions amplified from genomic clone pTA884, introducing external <i>KpnI</i> and <i>EcoRI</i> sites, used to ligate into pTA131, and internal <i>NdeI</i> site.

---

pTA1284	<i>rpa3ap</i> deletion construct pTA1282 with <i>trpA</i> marker digested from pTA1166 using <i>NdeI</i> and inserted at <i>NdeI</i> site in pTA1282, generating <i>trpA</i> -marked <i>rpa3ap</i> deletion construct.
pTA1285	<i>rpa3ap</i> deletion construct pTA1282 with <i>trpA</i> marker digested from pTA1166 using <i>NdeI</i> and inserted at <i>NdeI</i> site in pTA1282, generating <i>trpA</i> -marked <i>rpa3ap</i> deletion construct.
pTA1292	pGB68 marked with <i>hdrB</i> and promoter from pTA1185 at <i>BamHI</i> site.
pTA1293	pGB68 marked with <i>hdrB</i> and promoter from pTA1185 at <i>BamHI</i> site.
pTA1326	pTA1222 with <i>rpa1ap</i> , amplified from pTA937 introducing <i>BstEII</i> and <i>BamHI</i> sites, and inserted downstream of his-tagged <i>rpa1</i> at <i>BstEII</i> and <i>BamHI</i> sites.
pTA1327	Contains <i>rpa1</i> inserted upstream of His-tagged <i>rpa1ap</i> at <i>NdeI</i> site. <i>rpa1</i> was amplified from pTA937 introducing <i>NdeI</i> and <i>AseI</i> ( <i>NdeI</i> compatible) sites.

**Table 2.4 Oligonucleotides**

Oligoname	Sequence	Description	Used to construct
Hvo RPA1 F1	GCGGCTACACTTGGGACAGC	Forward PCR primer for <i>rpa1</i> probe	
Hvo RPA1 R1	CAACGGCGAGACGACAATCC	Reverse PCR primer for <i>rpa1</i> probe	
Hvo RPA2 F1	GCTCATCGCCACGAACTG	Forward PCR primer for <i>rpa2</i> probe	
Hvo RPA2 R1	GTGCGTTTCCCGTCTGCC	Reverse PCR primer for <i>rpa2</i> probe	
Hvo RPA3 F1	CCGCTCCCCGACTGGCTG	Forward PCR primer for <i>rpa3</i> probe	

Hvo RPA3 R1	ACAGACCCTGCTCAACGACAT C	Reverse PCR primer for <i>rpa3</i> probe	
ARPA1P1 F	CGCCGAGGGTTCGCTTC	Forward PCR primer for associated <i>rpa1</i> protein probe	
ARPA1P1 R	CGAATCCGAGCCTATCGCAG	Reverse PCR primer for associated <i>rpa1</i> protein probe	
ARPA1P2 F	GCCCGAGGTGTCGTTGAAC	Forward primer for <i>rpa1</i> associated phosphoesterase probe	
ARPA1P2 R	GTCGTCGCTGACTACCACGC	Reverse primer for <i>rpa1</i> associated phosphoesterase probe	
ARPA3P1 F	GCGGTCTCCGTCACCCAG	Forward primer for associated <i>rpa3</i> protein probe	
ARPA3P1 R	GCCCGTCTATCTGCTGCTCC	Reverse primer for associated <i>rpa3</i> protein probe	
Rpa2bamh1 F	CAGTAGGACGGATCCACCGCC GGTC	Downstream primer of <i>rpa2</i> containing a <i>Bam</i> HI restriction site	pTA931, pTA934
Rpa2kpn1 R	GCGTGGTACCCGAGTTTGACG GCGGC	Downstream primer for <i>rpa2</i> containing a <i>Kpn</i> I restriction site	pTA931, pTA934
Rpa2xba1 F	CGGGTCTAGATTCGCACGTGA ATCG	Upstream primer of <i>rpa2</i> containing a <i>Xba</i> I restriction site	pTA931, pTA934
Rpa2bamh1 R	TCCCGGATCCCGCCCATCAGG C	Upstream primer of <i>rpa2</i> containing a <i>Bam</i> HI restriction site	pTA931, pTA934
PEbamhI F	CGACGAGGATCCACCGTTTCG	Upstream primer of <i>rpa1</i> associated phosphoesterase containing a <i>Bam</i> HI restriction site	pTA940, pTA944

PEkpnI R	GCTCGGTACCGAGTCTGAGCC C	Upstream primer of <i>rpa1</i> associated phosphoesterase containing a <i>KpnI</i> restriction site	pTA940, pTA944
PEecorI F	GCCCGAATTCCGTCTGATTG	Downstream primer of <i>rpa1</i> associated phosphoesterase containing a <i>EcoRI</i> restriction site	pTA940, pTA944, pTA1189, pTA1191
PEbamHI R	CTCTCGCGGATCCTGCCCCG	Downstream primer of <i>rpa1</i> associated phosphoesterase containing a <i>BamHI</i> restriction site	pTA940, pTA944
Hvo RPA2bamI	GCGGTGGATCCGTCCTACTGC	PCR primer of <i>rpa2</i> containing <i>BamHI</i> restriction site	pTA930, pTA939, pTA938, pTA939
Hvo RPA2NdeI	AGGTGACGCCATATGGGCGTC ATCC	PCR primer of <i>rpa2</i> containing <i>NdeI</i> restriction site	pTA930, pTA939, pTA938, pTA939
Hvo RPA2NcoI	AGGTGACGCCCCATGGGCGTC ATCC	Over expression PCR primer of <i>rpa2</i> containing <i>NcoI</i> restriction site	pTA936, pTA939
Rpa1CF DS	GTTTCGAGGTACCGTTCGGGGA GC	<i>rpa1</i> deletion primers containing <i>KpnI</i> restriction site.	pTA1141, pTA1197
Rpa1CR DS	AGGTGCGCATATGAGCGCCTC GC	<i>rpa1</i> deletion primers containing <i>NdeI</i> restriction site.	pTA1141, pTA1197
trpANdeIF	CTCTGCACatATGTCGCTCGA AGACGC	Forward oligo containing <i>NdeI</i> restriction site to PCR amplify <i>trpA</i> and insert into a deletion construct	pTA1166, pTA1176

trpANdeIR	TGCATGCCatATgCGTTATGT GCG	Reverse oligo containing <i>NdeI</i> restriction site to PCR amplify <i>trpA</i> and insert into a deletion construct	pTA1166, pTA1176
Rpa3KpnI F	GCCGGTGGTACCACAGCCTC	Forward deletion oligo containing <i>KpnI</i> restriction site	pTA1142, pTA1198
Rpa3NdeIR	GCAAATCAGTCATATGCTACC TCGCC	Reverse deletion oligo containing <i>NdeI</i> restriction site	pTA1142, pTA1198
Rpa3EcoRIR	GACGGTGGAAATTCGGCCGTCG	Reverse deletion oligo containing <i>EcoRI</i> restriction site	pTA1142, pTA1198
Rpa3NdeI FC	GCGAGGTCGATGCATATGAGT TCCAACG	Oligo containing <i>NdeI</i> restriction site at <i>rpa3</i> ap start codon instead of <i>rpa3</i> stop codon.	pTA1142, pTA1198
Rpa1CF US	GGTCGAGTTCCATATGGTCGG GATTCGCC	Alternative forward <i>rpa1</i> deletion primer containing <i>NdeI</i> restriction site	pTA1141, pTA1197, pTA1189, pTA1191
Rpa1CR US	TACTACGTCTAGACGGACCTG TTCG	Alternative reverse primer for construction of <i>rpa1</i> deletion containing <i>XbaI</i> restriction site	pTA1141, pTA1197, pTA1189, pTA1191
Rpa3ap F DS	CTCGCTGAATTCGGTGGGTGC	<i>rpa3ap</i> forward deletion primer containing <i>EcoRI</i> restriction site	pTA1282
Rpa3ap R DS	CTGAGCGCATATGCGGGCGTC TCG	<i>rpa3ap</i> reverse deletion primer containing <i>NdeI</i> restriction site	pTA1282
Rpa1ap1 kpnI us	CCGCGAGTGGTACCGCAAGCC CG	Used to construct <i>rpa1ap</i> deletion plasmid.	pTA1170, pTA1171, pTA1217, pTA1219



Rpa1ap1 nsII us	CGACGACCGGCGATGCATTCA TGCGCGC	Used to construct <i>rpa1ap</i> deletion plasmid.	pTA1170, pTA1171
Rpa1ap1sphI ds	GCTGAAGGGCATGCGAGGCCG TGC	Deletion primer of downstream fragment of <i>rpa1ap1</i> containing <i>SphI</i> restriction site.	pTA1170, pTA1171
Rpa1ap1ecoR I ds	CGGCGAGAGAATTCCCTGCCC GGG	Deletion primer of downstream fragment of <i>rpa1ap1</i> containing <i>EcoRI</i> restriction site.	pTA1170, pTA1171
Rpa1F NcoI	CCCGACTCCATGGAACCTGAC C	Rpa1 forward primer containing <i>NcoI</i> restriction site	pTA1222
Rpa1apF BspHI	GGTGCCTCATGAGCGCCTCG	<i>rpa1ap</i> forward primer containing BspHI restriction site	pTA1223
Rpa1 CR US	TACTACGTCTAGACGGACCTG TTCG	Reverse oligonucleotide for deletion of <i>rpa1</i> containing a <i>XbaI</i> restriction site	pTA1141, pTA1197, pTA1189, pTA1191
Rpa1 CF US	GGTCGAGTTCCATATGGTCGG GATTCGCC	Forward oligonucleotide for deletion of <i>rpa1</i> containing a <i>NdeI</i> restriction site	pTA1141, pTA1197
RPEndel R DS	CTACCGGAACATATGACTCGG GTCG	Oligo for amplification of downstream region of <i>rpe</i> for <i>rpa1</i> operon deletion, containing internal <i>NdeI</i> site.	pTA1189, pTA1191
Rpa1ndeI F2	GTTGGACCCATATGTCGAACG ACG	Rpa1 oligonucleotide containing <i>NdeI</i> restriction site	pTA1189, pTA1191
Rpa3BspHI F	AGGTAGATCATGACTGATTTG C	Forward oligo for <i>rpa3</i> containing a <i>BspHI</i> restriction site	pTA1218

Rpa3APBspH I F	GTCGATGTTTCATGAGTTCCAA CG	Forward oligo for <i>rpa3ap</i> containing a <i>Bsp</i> HI restriction site	pTA1224
Rpa3apEcoRI R	CGGTCGGAATTCAGGCCGAC	Reverse oligo for <i>rpa3ap</i> containing an <i>Eco</i> RI restriction site	pTA1224, pTA1280
Rpa3ap FC US	CGTTGAACTCATATGCGAAGA CG	Forward oligo for <i>rpa3ap</i> containing <i>Nde</i> I site.	pTA1192
Rpa3 RAsel	CGAGTGGGGAATTCGTTGGAA TTAATTTACATC	Reverse primer for <i>rpa3</i> containing an AseI restriction site for ligation to <i>Nde</i> I end	pTA1218, pTA1281
Rpa1ap1SphI US	GCGATTTCCCGCATGCCGACG ACCG	Reverse oligonucleotide containing an <i>Sph</i> I site for creating deletion construct of <i>rpa1ap</i> not <i>trpA</i> <sup>+</sup> .	pTA1217, pTA1219
Rpa1apEcoRI HisR	CGTTCGGGGAATTCGCGCCTG C	Reverse oligo for inserting just <i>rpa1ap</i> into p963, contains <i>Eco</i> RI restriction site.	pTA1223
pyrE2F	CGGTATGGTCGTCTCTCTCG	Upstream primer of <i>pyrE2</i> for amplification, to check <i>rpa2</i> integration at <i>pyrE2</i> locus.	
Rpa2 F3	CACGCAGACCAAGACCGTCG	To check insertion of <i>Rpa2</i> at chromosomal <i>pyrE2</i> locus.	
Rpa1AseI/Ec oRI	CGGCGGCGAATTCGCGGTAGG CGATTAATCGCGTGC	Contains <i>Ase</i> I site to allow insertion of his-tagged <i>rpa1ap</i> downstream	pTA1222
Rpa1NdeI	CCCGACCATATGGAACCTCGAC C	Forward oligo allowing insert of non-histagged <i>rpa1</i> upstream of His- tagged <i>rpa1ap</i> in pTA1223.	pTA1327

Rpa1BstEI	CCGGCACGGTGACCGCCATCC	Allows insertion of <i>rpa1ap</i> downstream of his-tagged <i>rpa1</i> using existing <i>BstEI</i> site without disturbing operon structure.	pTA1326
Rpa3BstEI	GATGCGCGGTGACCTCGTGG	Allows insertion of <i>rpa3ap</i> downstream of his-tagged <i>rpa3</i> using existing <i>BstEI</i> site without disturbing operon structure	pTA1280
Rpa3NdeI	CGAGGTAGCATATGACTGATT TGCG	Forward oligo allowing insert of non-histagged <i>rpa3</i> upstream of His-tagged <i>rpa3ap</i> in p1224.	pTA1281
Rpa2F4	TATGAGTACGACGCTCGG	Check <i>rpa2</i> chromosomal integration.	pTA1282
Rpa2R4	GATGAGTCAGGTTCGACG	Check <i>rpa2</i> chromosomal integration.	
Rpa3ap gitR	TCGTTGGAcataTatgTTACATC GACCTCGC	Deletion oligo maintaining <i>rpa3</i> stop codon.	
Rpa3ap gitF	CTCCCAATGGgtACCAAGGTG GAGGC	US primer to insert whole of <i>rpa3</i> in deletion construct.	pTA1282
Rpa2RTF	GAAGTCGGCACCGTCCTCC	Diagnostic RT PCR primer to check <i>rpa2</i> chromosomal integration.	pTA1282
Rpa2RTR	GGTCCGCCCCGCTCGTCCC	Diagnostic RT PCR primer to check <i>rpa2</i> chromosomal integration.	
Rpa3opext F	GATGCTTTTCGAGGCGGTCTG	Oligo to test presence of <i>rpa3</i> operon in potential deletion.	

Rpa3opext R	CCTCGTTGACGCGAAACG	Oligo to test presence of <i>rpa3</i> operon in potential deletion.	
rpa1apBamHI	CGTTCGGGGGATCCGCGCCTG C	Reverse oligo of <i>rpa1ap</i> containing <i>Bam</i> HI site for insertion into p963	pTA1326, pTA1327
rpa1apEcoRI Cterm	CGGAGCCGGAATTGACGACG G	Oligo containing naturally occurring <i>Eco</i> RI site to insert C-terminus into p963 plasmids.	pTA1327
TrpAAS1	CGGAATCAGCGACGGCGACC	Internal <i>trpA</i> oligo to screen for potential deletions marked with <i>trpA</i> .	
Rpa3extF	CCGCGGGAGGGAGTGCGTCG	Oligo in external region of <i>rpa3</i> operon to screen for potential deletions.	p1179

### **2.1.2 *Haloferax volcanii* media**

Sterilised by autoclaving (1 minute 121°C) unless otherwise stated. Plates stored in sealed bags at 4°C and dried for at least 30 minutes before use.

**30% SW:** 240 g NaCl, 30 g MgCl<sub>2</sub>·6H<sub>2</sub>O, 35 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 7 g KCl, 20 ml 1 M Tris.HCl pH 7.5, dH<sub>2</sub>O to 1 litre (not autoclaved)

**18% SW:** 200 ml 30% SW, 100ml dH<sub>2</sub>O, 2 ml CaCl<sub>2</sub> added after autoclaving.

**Hv-Min salts:** 30 ml 1M NH<sub>4</sub>Cl, 36 ml 0.5 M CaCl<sub>2</sub>, 6 ml trace elements

**Hv-Min carbon source:** 41.7 ml 60% DL-lactic acid sodium salt (Sigma), 37.5 g succinic acid Na<sub>2</sub> salt.6H<sub>2</sub>O (Sigma), 3.15 ml 80% glycerol, dH<sub>2</sub>O to 250 ml, pH 6.5

**Trace elements:** 36 mg MnCl<sub>2</sub>·4H<sub>2</sub>O, 44 mg ZnSO<sub>4</sub>·7H<sub>2</sub>O, 230 mg FeSO<sub>4</sub>·7H<sub>2</sub>O, CuSO<sub>4</sub>·5H<sub>2</sub>O, dH<sub>2</sub>O to 100 ml

**10X YPC:** 50 g Yeast extract (Difco), 10 g Peptone (Oxoid), 10 g Casamino acids (Difco), 17.6 ml 1 M KOH, dH<sub>2</sub>O to 1 L (not autoclaved, used immediately)

**10X Ca:** 50 g Casamino acids, 23.5 ml 1M KOH, dH<sub>2</sub>O to 1 L (not autoclaved, used immediately)

**Hv-YPC agar:** 5 g Agar (Difco Bacto), 100 ml dH<sub>2</sub>O, 200ml 30% SW, 33 ml 10X YPC, 2 ml 0.5 M CaCl<sub>2</sub>. Microwaved without 10X YPC to dissolve agar. 10X YPC added, then autoclaved. CaCl<sub>2</sub> added prior to pouring.

**Hv-Ca agar:** 5 g Agar (Difco Bacto), 100 ml dH<sub>2</sub>O, 200 ml 30% SW, 33 ml 10X Ca, 2 ml 0.5 M CaCl<sub>2</sub>, 36 µg biotin (Sigma), 288 µg thiamine (Sigma). Microwaved without 10X Ca to dissolve agar. 10X Ca added, then autoclaved. CaCl<sub>2</sub>, thiamine and biotin added prior to pouring.

**Hv-YPC broth:** 100 ml dH<sub>2</sub>O, 200 ml 30% SW, 33 ml 10X YPC, 2 ml 0.5 M CaCl<sub>2</sub>. Autoclaved and CaCl<sub>2</sub> added when cool.

**Hv-Ca broth:** 150 ml 30% SW (autoclaved), 7.5 ml 1 M Tris.HCl pH 7.5, 6.5 ml Hv-Min carbon source, 3 ml Hv-Min Salts, 500 µl 0.5 M KPO<sub>4</sub> buffer (pH 7.5), 250 µl 36 µg biotin (Sigma), 288 µg thiamine (Sigma), 25 ml 10X Ca. All additives added once cooled.

**Table 2.5 *Haloferax volcanii* media supplements**

All media supplements were supplied by Sigma. All solutions were sterilised by filtration through a 0.2 µm filter (except for mevinolin).

Supplement	Abbreviation	Final concentration	Stock concentration
Leucine	+ Leu	50 µg/ml	10 mg/ml
Uracil	+Ura	50 µg/ml	50 mg/ml
Thymidine	+Thy	50 µg/ml (+50 µg/ml hypoxanthine in Hv-Ca)	4 mg/ml
Tryptophan	+Trp	50 µg/ml	10 mg/ml
5-FOA	+5-FOA	50 µg/ml (+10 µg/ml uracil)	50 mg/ml
Mevinolin	+Mev	4 µg/ml	10 mg/ml
Mitomycin C	+MMC	0.02 µg/ml	2 µg/ml
Hydrogen peroxide	+H <sub>2</sub> O <sub>2</sub>	2, 4 mM	30% (w/w)
Phleomycin	+Phleo	0.5, 1 and 2 mg/ml	50 mg/ml

### **2.1.3 *Haloferax volcanii* buffers and solutions**

All solutions were sterilised by filtration through a 0.22 µm filter.

**Buffered Spheroplasting Solution:** 1 M NaCl, 27 mM KCl, 50 mM Tris.HCl pH 8.5, 15% sucrose

**Unbuffered Spheroplasting Solution:** 1 M NaCl, 27 mM KCl, 15% sucrose, pH 7.5

**Spheroplast Dilution Solution:** 23% SW, 15% sucrose, 37.5 mM CaCl<sub>2</sub>

**Regeneration Solution:** 18% SW, 1X YPC, 15% sucrose, 30 mM CaCl<sub>2</sub>

**Transformant Dilution Solution:** 18% SW, 15% sucrose, 30 mM CaCl<sub>2</sub>

**ST Buffer:** 1 M NaCl, 20 mM Tris.HCl pH 7.5

**Lysis Solution:** 100 mM EDTA pH 8, 0.2% SDS

### **2.1.4 *Escherichia coli* media**

Sterilised by autoclaving.

**LB:** 10 g tryptone (Difco Bacto), 5 g yeast extract (Difco), 0.5 g NaCl, 10 g agar (Difco Bacto) when required for LB agar, dH<sub>2</sub>O to 1 L

**MU:** 10 g tryptone (Difco Bacto), 5 g yeast extract (Difco), 10 g NaCl, 10 g agar (Difco Bacto) when required for LB agar, dH<sub>2</sub>O to 1 L

**SOC:** 20 g tryptone (Difco Bacto), 5 g yeast extract (Difco), 0.58 g NaCl, 0.186 g KCl, 2.03 g MgCl<sub>2</sub>, 2.46 g MgSO<sub>4</sub>, dH<sub>2</sub>O to 1 L

**Table 2.6 *Escherichia coli* media supplements**

Supplement	Abbreviation	Final concentration	Stock concentration
Ampicillin	+Amp	50 µg/ml	4 mg/ml
Tetracycline	+Tet	3.5 µg/ml	1 mg/ml
Isopropyl-β-D-thiogalactoside <sup>a</sup>	IPTG	29 µg/ml	8 mg/ml
5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside <sup>a</sup>	X-Gal	36 µg/ml	10 mg/ml

<sup>a</sup>BlueTech solution (Mirador)

#### **2.1.5 Other buffers and solutions**

**TE:** 10 mM Tris.HCl pH 8.0, 1 mM EDTA

**TBE:** 89 mM Tris, 89 mM boric acid, 2 mM EDTA

**TAE:** 40 mM Tris.HCl, 20mM acetic acid, 1 mM EDTA

**Denaturing Solution:** 1.5 M NaCl, 0.5 N NaOH

**20X SSPE:** 525.9g NaCl, 82.8g NaH<sub>2</sub>PO<sub>4</sub>, 28.2g EDTA in 3 L dH<sub>2</sub>O, pH 7.4

**100X Denhardt's Solution:** 2% Ficoll 400, 2% PVP (poly-vinyl pyrrolidone) 360, 2% BSA (bovine serum albumin, Fraction V)

**Prehybridisation Solution:** 24 ml dH<sub>2</sub>O, 12 ml 20X SSPE, 2 ml 20% SDS, 2 ml 100X Denhardt's solution

**Hybridisation Solution:** 1.5 g dextran sulphate, 9 ml 20X SSPE, 1.5 ml 20% SDS, 18 ml dH<sub>2</sub>O

**Low Stringency Wash Solution:** 2X SSPE, 0.5% SDS

**High Stringency Wash Solution:** 0.2X SSPE, 0.5% SDS

**Neutralising Buffer:** 1.5 M NaCl, 0.5 M Tris.HCl, 1 mM EDTA

**Gel Loading Dye (5X):** 50 mM Tris.HCl, 100 mM EDTA, 15% Ficoll (w/v), 0.25% Bromophenol Blue (w/v), 0.25% Xylene Cyanol FF (w/v)

### **2.1.5 Protein purification buffers and solutions**

#### **50 mM Tryptophan**

1.02 g for 100 ml (in H<sub>2</sub>O), shake at 37°C to dissolve, store in dark 4°C.

#### **25 mM Tryptophan in 18% SW**

1.02 g for 200 ml (in 18% SW), shake at 37°C to dissolve, store in dark 4°C.

#### **5 M NaCl**

146.1 g for 500 ml (in H<sub>2</sub>O), filter sterilized using 0.22 µl filter,.

#### **1 M HEPES pH 7.5**

60 g for 250 ml (in H<sub>2</sub>O), pH equilibrated to 7.5 with NaOH, filter sterilized using 0.22 µl filter, (stored at 4°C).

#### **1 M Imidazole**

1.36 g for 20 ml (in H<sub>2</sub>O) or 3.4 g for 50 ml, filter sterilized using 0.22 µl filter (stored at 4°C).

#### **100 mM PMSF (phenylmethanesulfonyl fluoride)**

35 mg PMSF (Sigma) in 2 ml 100% ethanol (prepared fresh)

#### **Buffer A (protein purification)**

1 ml 20 mM HEPES pH 7.5, 20 ml 2 M NaCl, 500µl 1mM PMSF, imidazole to desired concentration, H<sub>2</sub>O to 50ml.

#### **72 % TCA (trichloroacetic acid)**

7.2 g TCA (Fisher) in 10 ml H<sub>2</sub>O (filter sterilised)

#### **0.15 % Deoxycholate**

0.015 g deoxycholic acid (Sigma) in 10 ml H<sub>2</sub>O (filter sterilised)

#### **Resuspension buffer**

2 ml 4 % SDS, 2 ml 0.2 M Tris pH 7.4, 1.5 ml 0.15 M NaOH, H<sub>2</sub>O to 10ml (filter sterilised).

#### **12.5% SDS-PAGE gel (resolving)**

3.13 ml 30% acrylamide/bis acrylamide (Severn Biotech Ltd), 1.88ml 1.5M Tris pH 8.8 (0.4% SDS), 37.5 µl 10% AMPS (ammonium persulfate, Sigma), 3.75 µl TEMED (tetramethylethylenediamine, Sigma), 2.46 ml H<sub>2</sub>O.



### **3.0% SDS-PAGE gel (stacking)**

0.77 ml 30% acryl/bis, 0.2ml 0.5M Tris pH 6.8 (0.4% SDS), 25 µl 10% AMPS (ammonium persulfate, Sigma), 2.5 µl TEMED (tetramethylethylenediamine, Sigma), 0.77 ml H<sub>2</sub>O.

### **SDS-PAGE running buffer**

0.25 M Tris, 1.92 M glycine, 1% SDS.

### **SDS PAGE sample buffer (4X)**

120 mM Tris.Hcl pH6.8, 8% SDS. 20% glycerol, 1 mg/ml bromophenol blue. Filter sterilised.

## **2.2 Methods**

### **2.2.1 Nucleic Acid Manipulation**

#### **Restriction Digests**

All restriction enzymes were supplied by New England BioLabs (NEB) and restriction digests were carried out following the manufacturer's guidelines and eluted in 30 µl. If no suitable buffer was available for a double digest single sequential digests were carried out where the first restriction digest was purified.

#### **PCR Amplification**

DNA polymerases DyNAzyme EXT and Phusion were supplied by Finnzymes. DyNAzyme EXT was used to amplify genomic DNA template and for diagnostic tests, Phusion was used for higher fidelity amplifications of DNA. For DyNAzyme EXT 1x Optimized DyNAzyme Buffer, 5% DMSO, 200 µM of each dNTP, 0.5 µM of each primer, 10-50 ng of template DNA and 1 U of DyNAzyme were used in each reaction. For Phusion 5x Phusion GC Buffer, 5% DMSO, 200 µM of each dNTP, 0.5 µM of each primer, 10 ng of template DNA and 1 U of Phusion were used in each reaction. For diagnostic tests 20 µl final volume reactions were used and 50 µl final volume reactions used for nucleic acid purification. Reaction conditions are shown in Table 2.7 and annealing temperatures were calculated using the basic method (Wallace, Shaffer *et al.* 1979). All reactions were carried using a Techne TC-512 thermocycler.

**Table 2.7 PCR reaction conditions**

Step	DyNAzyme	Phusion
Initial denaturation	92-94°C, 120 s	98°C, 30 s
Denaturation	94°C, 15 s	98°C, 10 s
Annealing	T <sub>m</sub> °C, 20 s	T <sub>m</sub> °C, 15s
Extension	72°C, 40 s/kb	72°C, 20 s/kb
Final Extension	72°C, 300 s	72°C, 300s

**Dephosphorylation**

Removal of vector 5' phosphate groups to prevent self-ligation was performed using Antarctic phosphatase (NEB). For each reaction 5 U of Antarctic phosphatase and 1 X Antarctic phosphatase buffer were used. Reactions were incubated at 37°C for 45 min and heat inactivated at 65°C for 10 min.

**Ligation**

All ligations were performed using T4 DNA ligase (NEB). For each reaction 400 U of T4 DNA ligase and 1x T4 DNA ligase buffer were used, incubated at 15°C for 16 hours. Ligations consisted of a molar ratio of 3:1 of insert to vector and were ethanol precipitated, resuspended in 4µl and transformed into *E. coli* or *H. volcanii*.

**Nucleic acid purification**

Restriction digest, PCR, desphosphorylated DNA products were all purified and extracted from agarose gels using NucleoSpin ExtractII kit (MN), where DNA is bound pH dependently to a silica membrane and was separated from contaminants by ethanolic washing. DNA was eluted in 30 µl buffer NE (5 mM Tris HCL pH 7.5).

For DNA ligations ethanol precipitation was used. For each purification 0.1 volumes of 3 M sodium acetate (pH 5.3) and 2 volumes of 100% ethanol were added to the DNA samples and incubated at -20°C for 1 hour. Samples were centrifuged at 4°C, 14,000 x g for 30 minutes. The supernatant was removed and the pellet was washed in 70% ethanol by centrifugation at 4°C, 14,000 x g for 15 minutes. The supernatant was removed and the pellet was air dried before resuspension in 4 µl of sterile H<sub>2</sub>O.

### **Nucleic acid quantification**

To determine the concentration and purity of plasmid preparations the 260:280 nm absorbance ratio of a spectrometer was used (Beckman Coulter DU 530)

### **Plasmid DNA extraction**

Plasmid DNA extraction from *E. coli* was performed using Macherey-Nagel (MN) NucleoSpin Plasmid kits following manufacturer's guidelines, using 2 ml of culture. For ultra pure plasmid DNA extractions, NucleoBond Xtra Midi kit (MN) using 333 ml of culture followed by an isopropanol and ethanol precipitation, plasmid DNA was resuspended in a final volume of 200 µl of TE. Ultra pure DNA plasmid samples were used for frozen stocks stored at -20°C and transforming into *H. volcanii*.

Extraction of PTA963 derivative plasmid DNA from *H. volcanii* H1209 was performed using 10 ml of culture, a 100 µl of ST buffer, 10 µl of EDTA and 150 µl of A1 buffer were used for resuspension after which the protocol *E. coli* plasmid extractions was followed.. For ultra pure plasmid DNA extractions from *H. volcanii* 150 ml of culture was used, 500 µl ST buffer, 20 µl EDTA and 5.5 ml of A1 buffer. After which the protocol for ultra pure *E. coli* plasmid extractions except for the extension to a 30 minute incubation with the A2 lysis buffer.

### **DNA sequencing and oligonucleotide synthesis**

All DNA sequencing reactions and analysis were performed by the Biopolymer Synthesis and Analysis Unit, University of Nottingham as was the majority of synthesis of oligonucleotides, Eurofins MWG Operon was used in the latter stages of the work.

#### **2.2.2 Gel Electrophoresis**

##### **Agarose gel**

All agarose gels were cast using AGTC Bioproducts Ltd hi-res standard agarose powder. TBE buffer was used for all gel electrophoresis except for 16 hour gels, which were used for Southern hybridisations, where TAE buffer was used. DNA samples were loaded with 1/5 final volume of 5X loading dye and electrophoresis was carried out at 5 V/cm. Ethidium bromide was added to diagnostic gels to a final concentration of 0.5 µg/ml. For gels where the DNA was to be extracted based on size, marker lanes were stained afterwards with ethidium bromide. A UV transilluminator was used to

visualize DNA and DNA samples were protected from UV damage by using foil whilst the correct DNA band was extracted.

### **2.2.3 Cloning and gene deletion analysis**

#### **Colony Lift**

*E. coli* and *H. volcanii* colonies and controls were picked using sterile sticks, patched onto plates and allowed to grow. Bio-Rad Zeta Probe positively charged membrane was left on the plate for 1 minute, transferred using forceps to Whatmann paper, and soaked in 10% SDS for 3 minutes. The filter was then placed on Whatmann paper soaked in denaturing solution for 5 minutes, then on Whatmann paper soaked in neutralisation solution for 3 minutes. Neutralisation was repeated and the filter was briefly washed in 2X SSPE before being air dried on Whatmann paper and UV-crosslinked at 120 mJ/cm<sup>2</sup>.

#### **Southern blotting by vacuum transfer**

A 200 ml 0.6% TAE agarose gel was run for 16 hours at 50V with buffer circulation for Southern analysis and then stained in 0.5 µg/ml ethidium bromide for 30 minutes. The gel-embedded DNA was then acid nicked with 0.25 M HCl for 20 minutes and washed for 10 minutes in H<sub>2</sub>O, followed by denaturing in denaturing solution for 45 minutes. A Bio-Rad Zeta Probe positively charged membrane, 15 cm x 25 cm, was washed in H<sub>2</sub>O and equilibrated in denaturing solution for 5 minutes. The gel was placed on top of the membrane on a Vacugene XL gel blotter and covered in denaturing solution and the transfer was run for 1 hour at 40 mbar using a Vacugene Pump (Pharmacia Biotech). The membrane was then washed in 2X SSPE and air-dried on Whatmann paper before being UV-crosslinked at 120mJ/cm<sup>2</sup>.

#### **Hybridisation**

Membranes from colony lifts and DNA transfers were incubated in 40 ml of prehybridisation solution and 800 µl of denatured fish DNA (10 mg/ml Roche) for 3 hours at 65°C. The prehybridisation solution was then replaced with 30 ml of hybridisation solution before addition of radiolabelled DNA probe and hybridised overnight at 65°C. The membrane was then washed twice in 50 ml of 2 X SSPE 0.5% SDS first for 10 minutes and then for 30 minutes. The membrane was then washed twice in 50 ml of 0.2 X SSPE 0.5% SDS each for 30 minutes. After which the

membrane was air dried on Whatman paper, wrapped in Saran wrap and exposed using a phosphorimager screen (Fujifilm BAS Cassette2 2325) overnight minimum. A Molecular Dynamics STORM scanner was then used to scan the screen.

### **Creating a radiolabelled DNA probe**

10 µl of H<sub>2</sub>O, 3 µl of template and 2 µl of ladder, if required, were boiled for 5 minutes at 100°C and then chilled on ice before incubation with 4 µl of High Prime random priming mix (Roche) and 0.37 Mbq of  $\alpha$ -<sup>32</sup>P dCTP (Perkin Elmer) at 37°C for 30 minutes. The DNA probe was purified using a BioRad P30 column, added to 450 µl of fish DNA (10 mg/ml Roche) and denatured by boiling.

### **Bioinformatics**

All primer design and sequence analysis was performed using MacVector 11.1 (MacVector Inc.). Sequence alignments were carried out using CLUSTALW and phylogenetic analysis using Neighbour Joining method (Hartman, Norais *et al.* 2010).

## **2.2.4 Escherichia coli Microbiology**

### ***E. coli* growth and storage**

All *E. coli* cultures were grown at 37°C for overnight shaking at 60 rpm in a floor shaker (or rotating at 8 rpm in a static incubator (LEEC). Cultures on solid media were incubated overnight at 37°C in a static incubator. Both liquid and solid media cultures were stored at 4°C short-term. For long-term storage, 25% (v/v) glycerol was added to cultures before they were snap frozen using dry ice and stored at -80°C. Ampicillin (100 µg/ml) was used for selection in both LB agar and broth. The optical density (O.D.) of cultures was measured at A<sub>650</sub> using a spectrometer.

### **Preparation of electrocompetent cells**

An overnight of culture of *E. coli* XL-1 or N2338 was grown in LB broth, XL-1 with tetracycline (3.5 µg/ml), then diluted 1/100 in LB broth with the appropriate antibiotic selection and incubated at 37°C until reached an A<sub>650</sub>= 0.5 – 0.8. The cultures were centrifuged at 6000 x g for 12 minutes at 4°C and resuspended in an equal volume of ice-cold sterile 1 mM HEPES (pH 7.5). This process was repeated resuspending in 0.5 volumes of ice-cold sterile 1 mM HEPES (pH 7.5), 0.25 volumes of ice-cold sterile 1 mM HEPES (pH 7.5), 0.1 volumes of ice-cold sterile 1 mM HEPES (pH 7.5) + 10%

glycerol and 0.001 volumes of ice-cold sterile 1 mM HEPES (pH 7.5) + 10% glycerol. Cells were snap frozen using dry ice before storage at -80°C.

### **Electroporation transformation of *E. coli***

1 µg of DNA in 4 µl of sterile H<sub>2</sub>O was added to 40 µl of electrocompetent on ice and pipetted into a pre-chilled electroporation cuvette (1 mm electrode gap GENEFLOW). The cuvette was pulsed at 1.8 kV in an *E. coli* gene pulser (BioRad) and immediately SOC was added. After a 40 minute incubation at 37°C with 8 rpm rotation the culture was plated on Mu agar with ampicillin selection.

### **Generation of unmethylated (*dam*<sup>-</sup>) plasmid DNA**

*H. volcanii* possesses restriction endonucleases that cleave DNA at 5'-GA<sup>(CH<sub>3</sub>)</sup>TC-3' sequences, therefore to transform plasmid DNA into *H. volcanii* unmethylated DNA was required. 10 ng of plasmid DNA extracted from XL-1 was transformed into *E. coli dam*<sup>-</sup> strain N2338, plated out on Mu and ampicillin plates and incubated for overnight at 37°C. Unmethylated DNA was then extracted as usual.

## **2.2.5 Haloferax volcanii microbiology**

### ***H. volcanii* growth and storage**

*H. volcanii* cultures were grown at 45°C rotating at 8 rpm in a static incubator. Cultures on solid media were grown at 45°C in the same incubator in a plastic bag to prevent desiccation. Liquid cultures and solid media cultures were stored at room temperature short-term. For long-term storage 25% (v/v) glycerol was added to cultures before they were snap frozen using dry ice and stored at -80°C.

### ***H. volcanii* genomic DNA preparation**

A 10 ml culture of *H. volcanii* was grown overnight at 45°C until at an O.D.<sub>650</sub> 0.6 - 0.8 and centrifuged at 3300 x g, 25°C for 5 minutes. The supernatant was removed, the cells resuspended in 200 µl of ST buffer and mixed by inversion with 200 µl of lysis buffer. The sample was overlaid with 1 ml of ethanol (100%) and DNA was spooled onto the interface of a gel-loading tip until solution was homogenously clear. The spool of DNA was then washed twice in ethanol (100%) and left to air dry. Samples were resuspended in 500 µl of TE, 50 µl of 3M sodium acetate (pH 5.2) and 400 µl of isopropanol was added and inverted to mix. Samples were then centrifuged at 3300 x g, 25°C for 5 minutes, supernatant removed and the pellet washed in ethanol

(70%). When pellets were thoroughly dry they were resuspended in 100  $\mu$ l TE at 45°C, shaking at 450 rpm for at least one hour followed by resuspension over night at 4°C. All genomic DNA samples were stored at 4°C.

### Obtaining a genomic clone

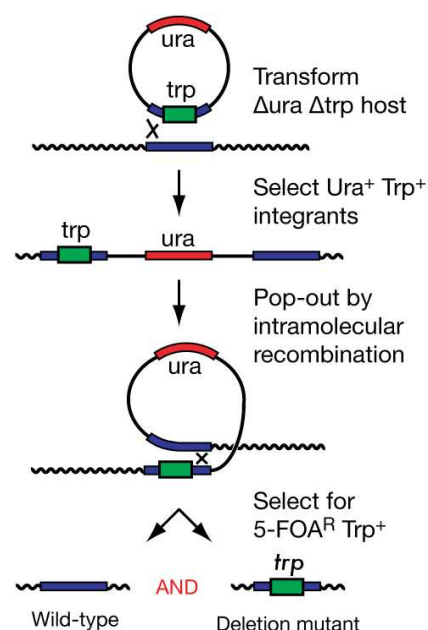
A genomic clone of the gene was isolated by size-selected genomic restriction digest, where sites at least 1 kb upstream and downstream of the gene of interest were chosen. The digested genomic DNA was electrophoresed on an agarose gel at 50 V overnight. The band of expected size (plus and minus 0.5 Kb) was extracted, the DNA purified and ligated to compatible sites in pBluescript. The resulting plasmid was transformed into *E. coli* and colonies screened by colony hybridisation using a gene-specific probe.

### Creating a genomic deletion construct

A deletion construct was made using the isolated genomic clone. Flanking regions of ~500 bp of the gene of interest were amplified by PCR with compatible internal sites and pTA131-compatible external restriction sites. The flanking regions were ligated together using internal restriction sites and then inserted into pTA131 using external restriction sites. A selectable autotrophic marker *hdrB* or *trpA* can be inserted in the deletion construct using internal restriction sites to bias generation of genomic deletion. The deletion construct was passed through the *dam*<sup>-</sup> *E. coli* N2338 strain and then transformed into *H. volcanii* to create a pop-in.

### Fig 2.1 Counter-selective pop-in/pop-out method

A *ura*<sup>-</sup> *trp*<sup>-</sup> host strain is transformed with a plasmid containing a *pyrE2* (*ura*) marker and a deletion construct of desired gene marked with tryptophan (Trp). The plasmid integrates at locus of gene to be deleted, conferring uracil and tryptophan prototrophy (*ura*<sup>+</sup> *trp*<sup>+</sup>). Selection for uracil is relieved and the plasmid recombines from the chromosome resulting in *ura*<sup>-</sup> derivatives that are selected for by plating out on 5-FOA. Selection is biased in favour of deletion mutant by selection for Trp<sup>+</sup>.



### ***H. volcanii* transformation**

A 10 ml culture of *H. volcanii* was grown overnight at 45°C until at an O.D.<sub>650</sub> 0.6 - 0.8 and centrifuged at 3300 x g for 8 minutes at 25°C in a round bottomed tube. The cells were gently resuspended in 2 ml of buffered spheroplasting solution and centrifugation repeated followed by resuspension in 600 µl of buffered spheroplasting solution. 200 µl aliquots per transformation were transferred to a new round bottomed tube and a 20 µl drop of 0.5M EDTA pH 8.0 was pipetted onto the side of the tube, gently inverted and incubated for 10 minutes at room temperature. The DNA sample made up of 10 µl *dam*<sup>-</sup> DNA (1-2 µg), 15 µl unbuffered spheroplasting solution and 5 µl 0.5M EDTA pH 8.0 was then added to cells in the same manner as the EDTA and incubated for 5 minutes at room temperature. 250 µl of 60% PEG solution, made up using unbuffered spheroplast solution, was then added and mixed by gently shaking horizontally ~10 times followed by incubation for 30 minutes at room temperature. 1.5 ml spheroplast dilution solution was then added, mixed by inversion and after 2 minutes incubation at room temperature was centrifuged at 3300 x g for 8 minutes at 25°C. The cell pellet was transferred whole in 1 ml regeneration solution (+ thymidine if required) to a sterile 4 ml tube. After an undisturbed recovery at 45°C for 90 minutes the pellet was resuspended by tapping the tube and incubated for a further 3 hours at 45°C rotating at 8 rpm. Cells were transferred to a round bottom tube and centrifuged at 3300 x g, 25°C for 8 minutes and resuspended gently in 1 ml transformant dilution solution. Serial dilutions were made using transformant dilution solution and 100 µl plated out on selective media.

### **Determining orientation of pop-in**

To determine the orientation of the pop-in, a genomic DNA preparation was made and restriction digested, preferably cutting once in the plasmid and in one of the regions of flanking homology, and analysed by Southern hybridisation.

### **Generating genomic deletion in *H. volcanii***

The pop-in was grown non-selectively in a 5 ml YPC culture (+thy) overnight at 45°C until at an O.D.<sub>650</sub> 0.6 - 0.8 and diluted 1/500 into a fresh culture. Relieving uracil selection allowed recombination between homologous flanking regions to occur, resulting in either a wild-type or a genomic deletion. This process was repeated twice and 100 µl of the resulting culture was plated out on Hv-Ca +5FOA at dilutions 10<sup>-1</sup>



and  $10^{-3}$ ; if autotrophic markers *hdrB* or *trpA* were used in the deletion construct these were selected for. As a positive control, a 100  $\mu$ l at dilution  $10^{-6}$  was plated on Hv-Ca +ura (+thy and/or +trp). By plating out on Hv-Ca +5FOA only cells which are *ura*<sup>-</sup> and therefore have lost the plasmid are selected, by selecting for an autotrophic marker only the cells that have performed recombination between homologous flanking regions that have generated a genomic deletion will be able to grow.

Resulting colonies (pop-outs) were patched out on selective media and grown for 2/3 days at 45°C and screened for deletion by colony hybridisation using a gene specific probe. Patches that did not hybridise to the probe were selected and genomic DNA preparations were made to confirm genomic deletion by restriction digest, using sites upstream and downstream of the deleted gene. Southern hybridisation was then performed using flanking regions in deletion construct as a probe.

#### **RNA extraction from *H. volcanii***

All equipment was wiped down with 0.1 M NaOH 1mM EDTA followed by 0.1% DEPC-treated water. Filter tips, individually wrapped tubes and RNase-free chemicals were used and gloves were worn at all times.

A 5 ml culture of *H. volcanii* was grown overnight at 45°C until at an O.D.<sub>650</sub> ~ 0.5 and diluted into a fresh culture using a range of dilutions and grown overnight at 45°C until at an O.D.<sub>650</sub> ~ 0.5. This was then diluted into a fresh 10 ml culture and grown overnight at 45°C until desired O.D. was reached. The culture (amount depending on O.D.) was then centrifuged at 3300 x g for 8 minutes at 25°C in a round bottomed tube. The cells were resuspended in 250  $\mu$ l of unbuffered spheroplasting solution and transferred to a 1.5 ml tube. 500  $\mu$ l of Trizol LS was added and homogenized by vortexing, followed by a 5 minute incubation at room temperature. 250  $\mu$ l of chloroform was then added and samples were vortexed for 30 seconds before incubation for 3 minutes at room temperature. The samples were then centrifuged at 4°C, 14,000 x g for 10 minutes, the resulting top aqueous layer was transferred to a fresh 1.5 ml tube. 500  $\mu$ l isopropanol was added and incubated for 10 minutes at room temperature. The samples were then centrifuged at 4°C, 14,000 x g for 10 minutes, supernatant removed and the precipitated RNA washed in 1 ml of ethanol (75%), followed by centrifugation at 4°C, 14,000 x g for 10 minutes. The supernatant was

removed and the pellet left to air-dry before resuspension in 45 µl of H<sub>2</sub>O. A 5 µl aliquot was used in a 1/10 dilution with TE to check RNA concentration and purity by measuring 260:280 nm absorbance ratio of a spectrometer (Beckman Coulter DU 530). RNA samples were snap frozen using dry ice and stored at -80 °C.

#### **DNase treatment of RNA samples**

5 µl of 10X TURBO DNase buffer and 1 µl of TURBO DNase (Ambion) were added to 45 µl RNA sample and incubated at 37°C for 30 minutes. A further 1µl of TURBO DNase was added and incubation repeated. 5 µl of DNase inactivation reagent (Ambion) was added, mixed by vortexing and incubated for 3 minutes at room temperature, vortexing occasionally. The samples were then centrifuged at 14,000 x g at room temperature and supernatant containing RNA transferred to a fresh tube. RNA samples were pipetted into 5 µl aliquots and snap frozen before being stored at -80°C. Filter tips, individually wrapped Eppendorf tubes and RNase-free chemicals were used and gloves were worn at all times.

#### **Reverse transcriptase PCR**

To amplify genomic RNA, OneStep RT-PCR Enzyme Mix (QIAGEN) was used comprising Omniscript Reverse Transcriptase, Sensiscript Reverse Transcriptase and HotStarTaq DNA polymerase (all QIAGEN). In one 20 µl reaction, 2 µl of OneStep RT-PCR Enzyme Mix, 1x OneStep RT-PCR buffer (QIAGEN), 10 mM of each dNTP, 1x Q solution (QIAGEN, aids transcription of GC rich template), 5-10 units of SUPERnase-In (Ambion), ~100 ng RNA template and 0.6 µM of each primer were used. Reaction conditions are shown in Table 2.8 and annealing temperatures were calculated using nearest neighbour method. All reactions were carried using a Techne TC-512 thermocycler. A “no RT” control was used for all RT PCRs. Filter tips, individually wrapped Eppendorf tubes, RNase free PCR tubes and RNase-free chemicals were used and gloves were worn at all times.

**Table 2.8** RT PCR reaction conditions

Step	DyNAzyme
Reverse Transcriptase	50°C, 30 min
Initial PCR activation step	95°C, 15 min
Denaturation	94°C, 0.5 min
Annealing	T <sub>m</sub> °C, 0.5 min
Extension	72°C, 0.5 min/kb
Final Extension	72°C, 10 min

### **2.2.6 Phenotype analysis**

#### ***H. volcanii* UV sensitivity test**

A 5 ml culture of *H. volcanii* in Hv-YPC broth (+thy) was grown overnight at 45°C until at an O.D.<sub>650</sub> ~ 0.4, diluted into fresh Hv-YPC broth (+thy) using a range of dilutions and grown overnight at 45°C until at an O.D.<sub>650</sub> ~ 0.4. Serial dilutions of the cells were made using 18% SW of 10<sup>-1</sup> to 10<sup>-8</sup> and duplicate 20 µl drops of the appropriate dilutions were pipetted onto Hv-YPC (+thy). Once the spots had air-dried at room temperature, the plates were exposed to UV light at the rate of 1 J/m<sup>2</sup>/sec (254 nm peak). All UV exposure was carried out with room lights turned off, shielding from light using an additional black bag to prevent photo-reactivation with DNA repair, and incubated for 6 days at 45 °C.

#### ***H. volcanii* mitomycin C sensitivity test**

A 5 ml culture of *H. volcanii* in Hv-YPC both (+thy) was grown overnight at 45°C until at an O.D.<sub>650</sub> ~ 0.4 and diluted into a fresh Hv-YPC broth (+thy) culture and grown overnight at 45°C until at an O.D.<sub>650</sub> ~ 0.4. Serial dilutions of the cells were made using 18% SW of 10<sup>-1</sup> to 10<sup>-8</sup> and duplicate 20 µl drops of each dilution were pipetted onto Hv-YPC + mitomycin C (0.02 µg/ml, +thy). This concentration of mitomycin C was determined empirically. Once spots had air-dried at room temperature they were incubated at 45°C for seven days.

#### ***H. volcanii* H<sub>2</sub>O<sub>2</sub> sensitivity test**

A 5 ml culture of *H. volcanii* in Hv-YPC broth (+thy) was grown overnight at 45°C until at an O.D.<sub>650</sub> ~ 0.4 and diluted into a fresh Hv-YPC broth (+thy) culture and

grown overnight at 45°C until at an O.D. <sub>650</sub> ~ 0.4. 1 ml of culture was aliquoted into 1.5 ml screw cap tubes, 10 µl of H<sub>2</sub>O<sub>2</sub> diluted using 18% SW (or 10 µl of 18% SW as a control) was added to give a final concentration of 0, 2, 4 and 8 mM. After an incubation of 1 hour at 45°C shaking at 450 rpm using an Eppendorf thermomixer, serial dilution of the cells were made of 10<sup>-1</sup> to 10<sup>-8</sup> using 18% SW. Duplicate 20 µl drops were pipetted on Hv-YPC agar (+thy) and the spots were left to air-dry at room temperature before incubation at 45°C until colonies were visible.

### Competition assay

10 ml Hv-YPC +thy broth was inoculated with 1 colony of a wild-type *bgaHa*<sup>+</sup> strain and another culture with 1 colony of a mutant *bgaHa*<sup>-</sup> strain. The cultures were grown overnight at 45°C until at a OD<sub>650</sub> ≈ 0.4 (~10<sup>8</sup> cells), serial dilutions 10<sup>-1</sup> to 10<sup>-8</sup> of each culture were made using 18% SW and dilutions and spot tested out on Hv-YPC +thy. A 10 ml Hv-YPC +thy culture was inoculated with 10<sup>3</sup> wild-type cells and 10<sup>4</sup> mutant cells. The mixed culture was grown until OD<sub>650</sub> ≈ 0.4, serial dilutions 10<sup>-1</sup> to 10<sup>-6</sup> were made and 10<sup>-4</sup>, 10<sup>-5</sup> and 10<sup>-6</sup> were plated out on Hv-YPC +thy. A fresh 10 ml Hv-YPC +thy culture was inoculated with 10<sup>4</sup> mixed cells, this was repeated three times. Once colonies were visible they were sprayed with an X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) solution (BlueTech, Mirador) and incubated overnight at 45°C.

The data was then used to calculate the growth defect of the mutant strain compared to the wild-type, using the following formula:

$$\text{Growth defect (\%)} = \left( 1 - \left( \frac{\text{fraction of mutant cells at the end}}{\text{fraction of mutant cells at the start}} \right)^{-G} \right) \times 100$$

Where G = the number of generations

During every generation the number of cells doubles therefore 2<sup>G</sup> = 10<sup>n</sup> and thus the cells were diluted 10<sup>n</sup>-fold.

$$G = \frac{n \ln 10}{\ln 2}$$

### **2.2.7 Protein purification in *Haloferax volcanii***

#### **Ni<sup>2+</sup> charging of pull-down column**

IMAC Sepharose 6 Fast Flow beads (GE Healthcare) were charged with Ni<sup>2+</sup> as follows. 0.5 ml was used per column. The beads were washed twice with at least 2 volumes of dH<sub>2</sub>O, and then equilibrated for 30 minutes with at least 0.2 volumes of 0.2 M NiSO<sub>4</sub>. Two further washes with dH<sub>2</sub>O were performed, the beads were then washed once with 5 ml of Buffer A (containing 500mM imidazole) then a further three times with Buffer A (containing 20mM imidazole). Beads were resuspended in 0.5 ml Buffer A (20mM imidazole) per sample.

#### **Protein overexpression and purification**

A starter culture was grown overnight in 50 ml Hv-Ca broth to OD<sub>650</sub> of 0.6 and used to inoculate 343 ml Hv-YPC broth. The culture was incubated at 45°C with shaking (175 rpm) overnight to an OD<sub>650</sub> of 0.5, when protein expression was induced by adding 0.25g of 25 mM Trp dissolved in 18% salt water (SW), and the culture was incubated at 45°C with shaking for a further 1 h until OD<sub>650</sub> ≈ 0.7. The culture was then centrifuged at 3,300 x g for 8 min at 4°C.

Cells were resuspended in 7 ml ice-cold Buffer A (20 mM imidazole) and lysed by sonication (2-3x30 sec at ≤ 8μ) on ice until the suspension was no longer turbid. The cell lysate was then transferred to a 15ml round bottomed tube, and centrifuged at 16,000 x g for 15 min at 4°C. A 75 μl sample was taken, and the remainder incubated overnight at 4°C with 0.5 ml of IMAC Sepharose 6 FastFlow beads (GE Healthcare) that had been charged with Ni<sup>2+</sup> and equilibrated in binding buffer. The slurry was applied to a Poly-Prep column (Bio-Rad), and the flow-through was collected and reloaded onto the column, followed by three washes with 4 ml of ice-cold buffer A (20 mM imidazole) and one wash with 1 ml ice-cold buffer A containing 50 mM imidazole. Bound protein was eluted with 1 ml of buffer A containing 500 mM imidazole.

#### **SDS-Polyacrylamide Gel Electrophoresis**

Protein samples were mixed with a 1/4 (final volume) of 4x gel loading dye and run on 12.5% SDS-PAGE gels (with a 3.0% stacking gel) unless otherwise stated using a cassette (Invitrogen). First a loading gel was poured, with a layer of isopropanol on

top to ensure a flat surface and air lock the cassette. Once the resolving gel had set, the isopropanol was thoroughly washed off and the stacking gel was poured followed by the insertion of a comb. A PageRuler size ladder (Fermentas) was run alongside protein samples. Gels were run for ~1hour 20 minutes in 1x SDS-PAGE running buffer (200 V, 36 mA per gel). Proteins present were stained with colloidal coomassie blue (Invitrogen).

### **Protein precipitation**

To the protein sample, 1/10<sup>th</sup> of the volume of 0.15% deoxycholate was added and vortexed before incubating for 10 minutes at room temperature. This was followed by addition of 1/10<sup>th</sup> of the original volume of 72% trichloroacetic acid (TCA) and incubation at room temperature for 5 minutes. Samples were then centrifuged at 14 000 xg at room temperature for 8 minutes. The supernatant was removed and the precipitated protein resuspended in 15 µl resuspension buffer and 5 µl 4x SDS PAGE sample buffer. The samples were heated for 10 minutes at 94°C and cooled on ice before loading onto a SDS PAGE gel.

### **2.2.8 Protein analysis**

#### **Processing and Tryptic Digestions**

##### **Performed by Susan Liddell.**

Samples were prepared and separated by SDS-PAGE as described earlier, and stained using colloidal coomassie blue (Invitrogen). Samples were excised from gels using a sterile scalpel and processed in gel pieces using the ProteomeWorks MassPREP robotic liquid handling station (Waters, Elstree, UK). The samples were incubated three times, in 100 µl of de-stain solution (50 mM ammonium bicarbonate, 50% acetonitrile) for 10 minutes at room temperature. Following removal of the final aliquot, the sample was dehydrated in 50 µl of acetonitrile and incubated for 5 minutes at room temperature. The acetonitrile was removed and the gel plugs incubated for 10 minutes to allow evaporation. The sample was further processed by incubation at room temperature in 50 µl of reducing solution (10 mM dithiothreitol, 100 mM ammonium bicarbonate) for 30 minutes and, followed by removal of the reducing solution and incubation at room temperature in 50 µl of alkylation solution (55 mM iodoacetamide, 100 mM ammonium bicarbonate) for 20 minutes. The gel slice was then washed at room temperature in 50 µl of 100 mM ammonium bicarbonate for 10

minutes, 50  $\mu$ l of acetonitrile for 5 minutes and dehydrated by double room temperature washes in 50  $\mu$ l of acetonitrile for 5 minutes and evaporation for 5 minutes. The microtitre plate containing the gel plugs was cooled to 6°C for 10 minutes before addition of 25  $\mu$ l per well of trypsin gold (Promega), diluted to 10 ng  $\mu$ l<sup>-1</sup> in trypsin digestion buffer (50 mM ammonium bicarbonate). The plate was incubated at 6°C for a further 20 minutes to permit trypsin entry into the gel plugs with minimal autocatalysis before incubation at 40°C for 4 hours. Samples were stored at 4°C until MS analysis.

### **Mass Spectrometry, data dependant acquisitions (DDA)**

#### **Performed by Susan Liddell.**

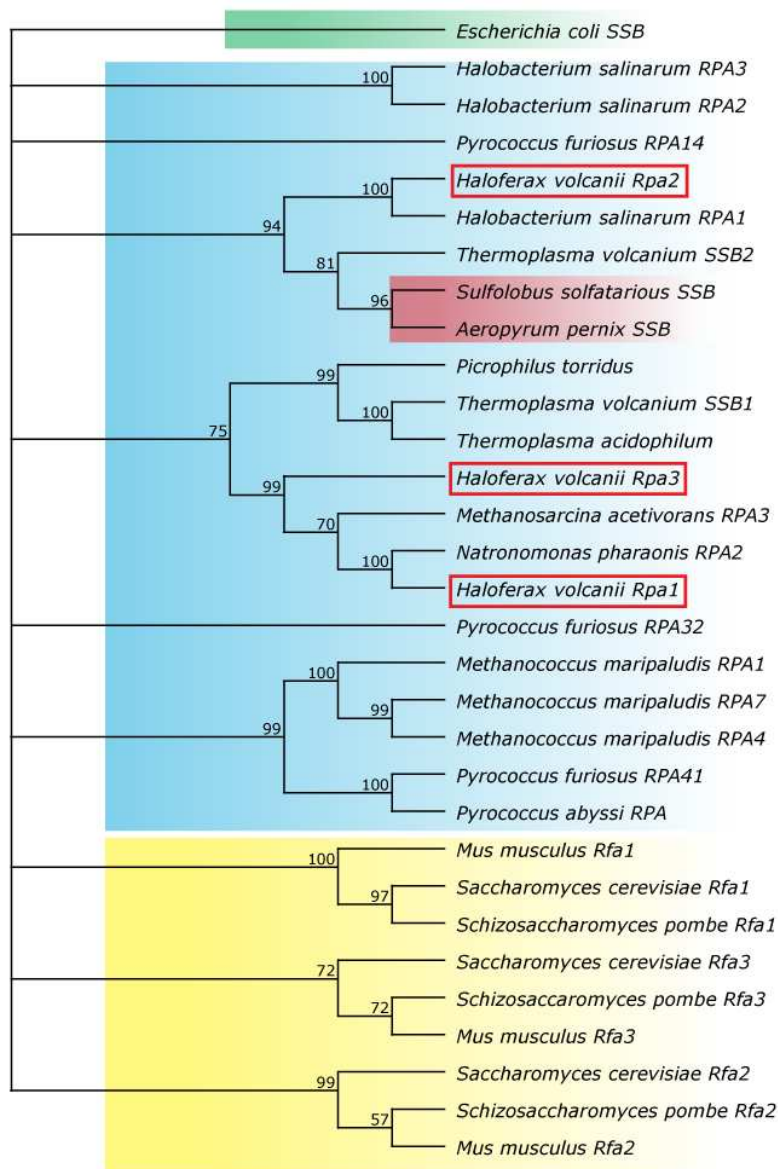
Samples were analysed by LC-ESI-tandem MS on a Q-TOFII mass spectrometer fitted with a nanoflow ESI (electrospray ionization) source (Waters Ltd). Peptides were separated on a PepMap C18 reverse phase, 75  $\mu$ m internal diameter, 15-cm column (LC Packings), involving the use of a short-precursor (C18) where the peptides were trapped and desalted. The peptides were then eluted over a short solvent gradient of 10 minutes, before mass spectrometry analysis. The mass spectrometer was operated with a capillary voltage of 3000 V in positive ion mode, using argon as the collision gas. Tandem MS data was acquired using an automated data-dependent switching between MS and MS/MS scanning based upon ion intensity, mass and charge state (data directed analysis (DDA<sup>TM</sup>)). In this automated acquisition type of experiment, a method was created in the MassLynx 4.0 software in which charge state recognition was used to select doubly, triply and quadruply-charged precursor peptide ions for fragmentation. One precursor mass at a time was chosen for tandem MS acquisition. A collision energy was automatically selected based on charge and mass of each precursor and varied from 15 to 55 eV.

ProteinLynxGlobalServer version 2.0 (Waters, Ltd) was used to process the uninterpreted MS data into peak list (pkl) files which were searched against all entries in Swissprot and/or NCBI nr databases using the web version of the MASCOT MS/MS ions search tool (<http://www.matrixscience.com/>). Carbamidomethylation of cysteine and oxidation of methionine were set as variable modifications. One missed cleavage by trypsin was accepted. Other than file type (Micromass pkl) and instrument type (ESI-QUAD-TOF), all remaining search values were the pre-set defaults.

## Chapter 3: Replication protein A of *Haloferax volcanii*

### 3.1 Phylogenetic analysis of archaeal RPA

*Haloferax volcanii* possess three replication protein A (RPA) genes *rpa1*, *rpa2*, and *rpa3*, which were identified in the *H. volcanii* genome that had been sequenced by The Institute for Genomic Research (TIGR) (Hartman, Norais *et al.*). *Pyrococcus furiosus* also possesses three RPAs of 41 kDa, 32 kDa and 14 kDa in size, which have been shown to function as a heterotrimer. However phylogenetic analysis of archaea RPA subunits, shown in Figure 3.1, shows that the *P. furiosus* RPA subunits do not fall into uniformed clades, as seen with the eukaryotic RPA subunits, which function as a heterotrimer. Therefore, the *P. furiosus* subunits functioning as a heterotrimer is likely to be an exception amongst archaea.



**Figure 3.1.**  
**Phylogenetic analysis of *H. volcanii* RPAs**

Sequences were aligned using CLUSTALW (Gonnet matrix), phylogenetic analysis was generated using Neighbour Joining method with a 10000 bootstrap iterations, *E. coli* SSB was used to root the tree (Hartman, Norais *et al.* 2010). Red boxes outline RPA1, 2 and 3. *Methanococcus maripaludis* C5 strain was used.



The phylogenetic analysis shown in Figure 3.1 demonstrates that the number of RPAs present varies enormously from archaeal species to species, and where there are three RPAs as seen in *P. furiosus* they do not fall into distinct clades like the eukaryotic RPA subunits. The clustering of *H. volcanii* RPA1 and RPA3 together in the same clade suggests they are closely related and that one of them has most likely arisen through a recent gene duplication event. Therefore the ancestor of *H. volcanii* originally had two RPA subunits. RPA2 is more closely related to *Methanococcus maripaludis* RPAs than to *H. volcanii* RPA1 and RPA3, which suggests a divergence between the three RPAs, perhaps through structural adaptation for different cellular roles. In short, the phylogenetic analysis does not support the hypothesis that the *H. volcanii* RPA subunits form a heterotrimer, as seen in eukaryotes and *P. furiosus*.

This enormous variety in the number of archaeal RPA subunits and the quaternary structures they might form is seen also for the MCM subunits (Chia, Cann *et al.* 2010) (Walters and Chong 2010). A vast range of MCM subunits are seen, which differ from species to species and do not fall into highly-conserved, uniformed clades as seen for the six eukaryotic MCM subunits. The archaeal MCM subunits have presumably arisen from gene duplication events; this may have allowed the archaeal MCM subunits to become specialized in their cellular role. A similar situation is seen for RPA. The eukaryotic RPA subunits fall into three highly-conserved clades but the archaeal RPA subunits do not. This is even seen for the *P. furiosus* RPA subunits, which act as a heterotrimer like the eukaryotic RPA, but do not fall into uniformed clades, as shown in Figure 3.1. Therefore it is possible that the existence of three RPAs in *H. volcanii* is to enable the RPAs to have different, specialized roles in DNA replication and/or DNA repair. However it is unlikely they would achieve this as a heterotrimeric complex.

Both *rpa1* and *rpa3* are in operons with other genes; *rpa1* is in an operon with an OB-fold containing gene (hereby designated associated protein, or AP) and a calcineurin-like phosphoesterase, while only one OB-fold containing (AP) gene is present in the *rpa3* operon (see Figure 3.3). The presence of an associated protein gene featuring an OB-fold in the same operon as *rpa* can be found in other euryarchaeota, including *Halobacterium marismortui*, *Halobacterium salinarum*, and *Natronomonas pharaonis*, as well as in *Methanosarcina mazei* and *Methanosarcina barkeri*. The associated

protein gene has been assigned to the cluster of orthologous groups (COG) 3390 (Berthon, Cortez *et al.* 2008). An alignment of *H. volcanii* RPA1, 2, 3, RPA1AP and RPA3AP protein sequences is shown in Figure 3.2. The presence of a phosphoesterase in an *rpa* operon has never been reported until now.

```

Hvo Rpa1  - - - - - 10 20 30
Hvo Rpa2  - - - - -
Hvo Rpa3  - - - - -
Hvo rpa3ap - - - - -
Hvo rpa1ap M S A S P V Y G T R R I A Y R V F A A E F D D A S L S T S D

Hvo Rpa1  - - - - - 40 50 60
Hvo Rpa2  - - - - -
Hvo Rpa3  - - - - -
Hvo rpa3ap - - - - -
Hvo rpa1ap S D E E R A P N Y V V T F T G A R L N R T F V A G V L T E V

Hvo Rpa1  - - - - - 70 80 90
Hvo Rpa2  - - - - - M E L D Q H A F E L A S A L
Hvo Rpa3  - - - - - M G V I R E V Y D D L D T D V E F E F E A A V
Hvo rpa3ap - - - - - M T D L R T H A E I A D Q F
Hvo rpa1ap E H V N D R V L R G R I A D P T G A F V T Y A G Q Y Q P E P

Hvo Rpa1  - - - - - 100 110 120
Hvo Rpa2  G - - - - V D K E E V K S D L Q N L L - Q Y S V P L D E A K
Hvo Rpa3  N D K V R Q M G G L A D E E T A A V L A H E R R D E V V N
Hvo rpa3ap - - - - - S A D H F E R E L C V Y T H F P V D E A R
Hvo rpa1ap M A Y L D A A T P P A F V S L A G K A R T Y E P D A D V V

Hvo Rpa1  - - - - - 130 140 150
Hvo Rpa2  Q S V R R K H G G S S G G S D G A P A T K R E V D I D P D
Hvo Rpa3  G T A D I E P M R D D E F F A K V V S G E R T F E R D
Hvo rpa3ap - - - - - R S V V N S Y - - - - -
Hvo rpa1ap Y S V R P E S V N T H D A D V R D R W E V S A E A T L R

Hvo Rpa1  - - - - - 160 170 180
Hvo Rpa2  G G N V S Y T V R V L T V G T R S I V - - - - - Y Q G D F Q
Hvo Rpa3  G - - - - - - - - - - - - - - - - - - - - - E D R D G
Hvo rpa3ap - - - - - - - - - - - - - - - - - - - - - - L D E A
Hvo rpa1ap R I A V F D E A L S M P Y R G D D L T R A L E A R G V D S T

Hvo Rpa1  - - - - - 190 200 210
Hvo Rpa2  T T R E G E L A D E S G V I S Y T A W Q D F G F E P G D S V
Hvo Rpa3  R Y V N I E V A D E T G T I R V S L W D E M A A G A K E N L
Hvo rpa3ap V L G G V E V V L A G - - - - - Q T L L S D I D E D - - -
Hvo rpa1ap L A A G V P R A I D H Y T T R E Y L E A L R E V A V Q A L

Hvo Rpa1  - - - - - 220 230 240
Hvo Rpa2  V T G N A G V R E W D G R P E L N I G A S S T V G V E S E T
Hvo Rpa3  E G G T I R I G G R F K D G T A - G E V S A S K V E D
Hvo rpa3ap - - - - - - - - - - - - - - - - - - - - - -
Hvo rpa1ap E L V A G D R D Q V D P I D V A F G D G D A V L G P L P E

Hvo Rpa1  - - - - - 250 260 270
Hvo Rpa2  V R T P Y D D R T G G F A D L I D L Q A G D R G T V V R
Hvo Rpa3  L D A E V R V Q V L D S T R V T D L A G L S D V N L A G K
Hvo rpa3ap - - - - - E Q W V D V R A K V V E L W E - - - - -
Hvo rpa1ap E D L E P A A S H P T R G E A D A G F V E D A G E F E N

Hvo Rpa1  - - - - - 280 290 300
Hvo Rpa2  V E V D S R T I D G R N G E T F I L S G - - - - - V Y A D E T G
Hvo Rpa3  I L D A G T V R T P D R D D G T I G R V S N L S V G D T G
Hvo rpa3ap - - - - - - - - - - - P R S E S I A Q W G - - - - - L L G D D S G
Hvo rpa1ap E S G S E P A A T I D S E P E F D D G A E S E A V E P E S E

Hvo Rpa1  - - - - - 310 320 330
Hvo Rpa2  R E P F T D W A P - - - - - R P D V E G A S L R L S D V Y V
Hvo Rpa3  R E R V T L W D E R A D L A E E L E P G Q S V E V Y D Q Y V
Hvo rpa3ap R L K T T L E T R E - - - - - L P E L E V L Q A G R C S E H
Hvo rpa1ap E R L V I L V E - - - - - R R A D A V E L R V Q L R
S E P I A E P D S E P E S A D P E P V A A E T T A A E S

Hvo Rpa1  - - - - - 340 350 360
Hvo Rpa2  R E F R G V P Q V N L S E F T T L D V L D D P V S V T D S A
Hvo Rpa3  R E R D G S L E H V G R R G A V E D E E E T V P E T
Hvo rpa3ap D E Y Q G V F S V A L N R P T S T E L D E E E V G D S
Hvo rpa1ap M P A V A R G R V D L G D F G A D G D V D A V V G - - -
Q P G A S E S E P A S E P S T T V D S E P E P A A E P E

Hvo Rpa1  - - - - - 370 380 390
Hvo Rpa2  P R L K I G E A V D A G - - - - - G M F V F T L R V N
Hvo Rpa3  T D I G S I E L G - - - - - - - - - - - Q T V R I A G G V
Hvo rpa3ap - - - - - - - - - - - - - - - - - T S V F G A L L
Hvo rpa1ap S E P A A E P D T S I P E F F E V Y T D S L D E S E P D F

Hvo Rpa1  - - - - - 400 410 420
Hvo Rpa2  L E V - - - - - - - - - - - - - - - - - - - - - - K D G S G L
Hvo Rpa3  I E A D G K R T T R D D G S E G Q V R N I R Y K D G T G D
Hvo rpa3ap A D L D - - - - - - - - - - - - - - - - - - - - - - E D V L G R
Hvo rpa1ap D D G A L G D F F D A P D A G A T D A G S L D D S G S G T

Hvo Rpa1  - - - - - 430 440 450
Hvo Rpa2  T E R C P - - - - - - - - - - - K S R V V Q N G Q C R Q D
Hvo Rpa3  I R V A L W G E K A D A D V D L A D Y V I T D A E I K E G
Hvo rpa3ap I R K C P F - - - - - - - - - - - G C E R V L Q A G R C S E H
Hvo rpa1ap L G D F D D G F D D P D P E A G D S T D S T D S A S A D A D

Hvo Rpa1  - - - - - 460 470 480
Hvo Rpa2  G F V D - - - - - - - - - - - G E D D M R V A I L D D G T G T V
Hvo Rpa3  W D T E - - - - - - - - - - - L E A S R G C S V A M D R A P
Hvo rpa3ap G S V E - - - - - - - - - - - G E F D L R T A A V V D D G D E V H
Hvo rpa1ap P R R V - - - - - - - - - - - D D A T P P V I V Y L A D V F I L G
D S V D F D G M Y E L D F D R A E L E S E F G T E F T S G

Hvo Rpa1  - - - - - 490 500 510
Hvo Rpa2  T T A T L D R D L T - - - - - - - - - - - E E T I G G T M A D M E
Hvo Rpa3  E G A A G T D A G G S A P T P P S D E G I G A F S G D Q S S
Hvo rpa3ap E V I F N R E M T - - - - - - - - - - - E E L T G I F L D E A K Q
Hvo rpa1ap Q R A D E - - - - - - - - - - - E D A V G A I S R R E Q Q
A D V D F A G E A D I D V P - D A D D E T L E Q L E D S A A

Hvo Rpa1  - - - - - 520 530 540
Hvo Rpa2  A A R E A M D K E V Y A D D I A D T L E V G R E Y R V R N L
Hvo Rpa3  D D T S A A V G G S S D A S A A E S T G E A V E F T G T V
Hvo rpa3ap M A M D A L D T I V E E M R D L V G Y I R V T G - P
Hvo rpa1ap T T G R L V V R - F E L V T G V V L C E D A P G D F
A A E S D F E P A T A A A D P F A V Y D S E F F P T P D A

Hvo Rpa1  - - - - - 550 560 570
Hvo Rpa2  S V D E Y G A N I E T D E F E E T D D D P A E R - - - - -
Hvo Rpa3  V O A G T T V T L D D G T Q T K E V D T D A D L G L G D E V
Hvo rpa3ap T T G E V T L A V E V E L R E R T A D - - - - -
Hvo rpa1ap S S G D E Y G T H - - - - - - - - - - - F S A E D I D L E S V V D A
S D G D E D E A D A D P D A D E S A E D I D L E S V V D A

Hvo Rpa1  - - - - - 580 590 600
Hvo Rpa2  - - - - - - - - - - - A L A L E L T E R A - - - - -
Hvo Rpa3  T V S T E T D D A I S A S V L T E H G A Q Q - - - - -
Hvo rpa3ap - - - - - - - - - - - A E E L E I K A V S M - - - - -
Hvo rpa1ap M D D L D D G D G A T R D E V A A V V D E H G A D F G A Y

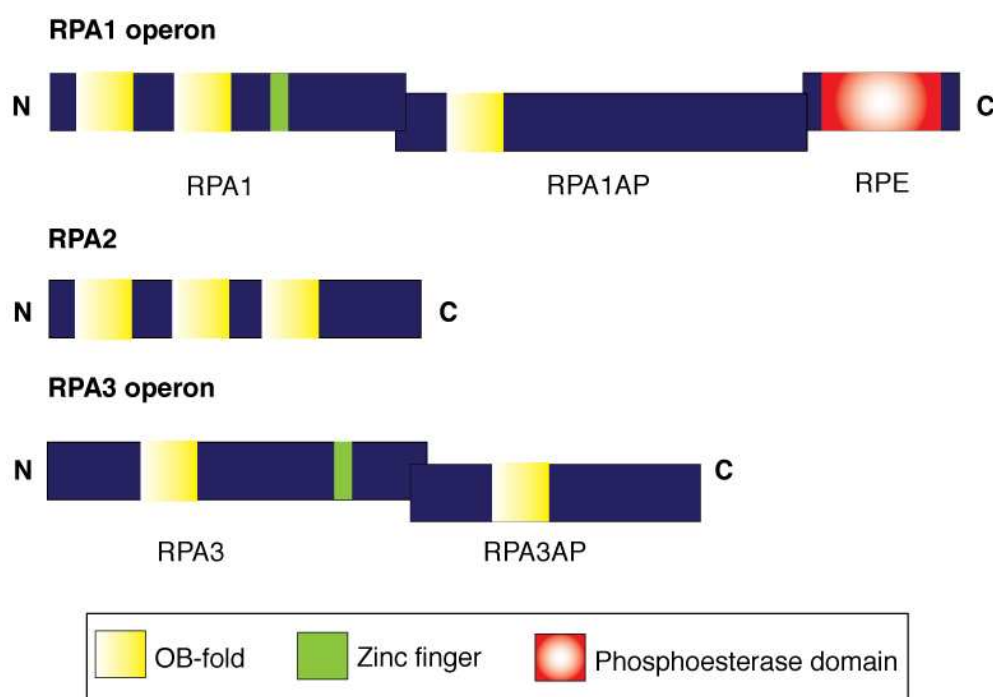
Hvo Rpa1  - - - - - 610 620 630
Hvo Rpa2  - - - - -
Hvo Rpa3  - - - - -
Hvo rpa3ap - - - - -
Hvo rpa1ap E D A I Q E A L E G G R C Y E P D G V L K A I

```

**Figure 3.1 Alignment of *H. volcanii* RPAs and APs.**

Alignment of *H. volcanii* (Hvo) RPA 1, 2 and 3, RPA1AP and RPA3AP was performed using CLUSTLEW (Gonnet matrix).

The predicted protein domains show that RPA1 and RPA3 both possess zinc fingers, however RPA1 has three OB-folds compared to the single OB-fold present in RPA3. Both associated proteins RPA1AP and RPA3AP possess a single OB-fold suggesting a possible role in DNA binding. The RPA1 phosphoesterase (RPE) has a calcineurin-like phosphoesterase domain, the presence of a phosphoesterase in an *rpa* operon could imply a role in processing DNA during DNA repair.



**Figure 3.3. Operon and domain structure of *Haloferax volcanii* RPAs and RPA associated proteins**

The three RPAs are shown, RPA1 and 3 in operons with their associated proteins, the RPA1 operon also encodes the Rpa1 phosphoesterase (RPE). The domain structures comprising of OB-folds, zinc fingers and a phosphoesterase domain are shown.

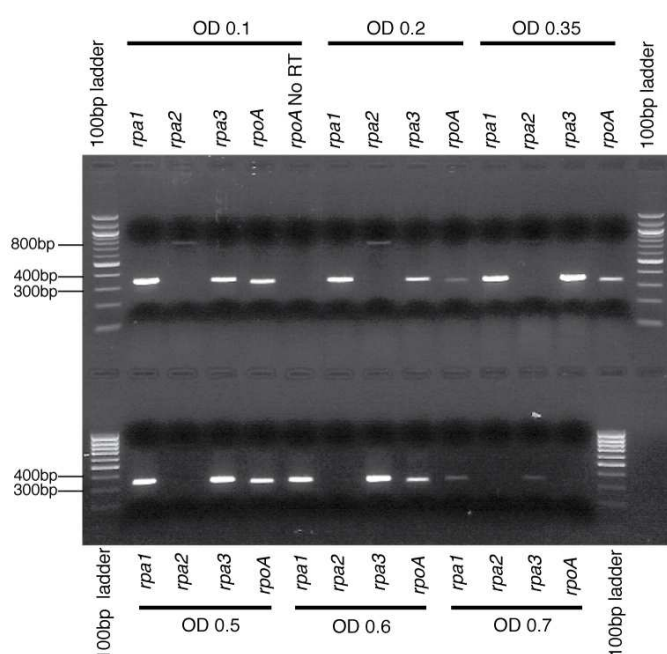
An alternative hypothesis to the three *H. volcanii* RPAs forming a heterotrimeric complex is that they form a complex with their respective associated proteins (APs). However as *rpa2* is not in an operon with an associated protein, instead RPA2 may form a homo-complex as seen in crenarchaea.

## 3.2 Expression of the three RPAs

To investigate whether all three RPA and their associated protein genes are expressed and if their expression differs during the cell cycle, which would suggest a specialized role in DNA replication, reverse transcriptase (RT) PCRs were carried out using RNA extracted from wild-type *H. volcanii* H678 at different points during the growth phase. The RT PCRs, shown in Figures 3.3 and 3.4, show that all three of the RPAs are expressed throughout the cell cycle.

**Table 3.1** *rpa1*, *rpa2* and *rpa3* internal RT primers

Primer	Target gene	Product size
Hvo RPA1 F1	<i>rpa1</i>	340 bp
Hvo RPA1 R1	<i>rpa1</i>	
Hvo RPA2 F1	<i>rpa2</i>	759 bp
Hvo RPA2 R1	<i>rpa2</i>	
Hvo RPA3 F1	<i>rpa3</i>	345 bp
Hvo RPA3 R1	<i>rpa3</i>	

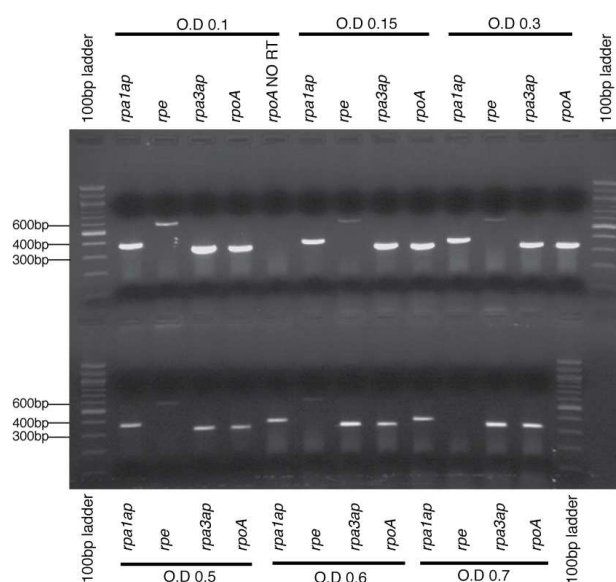


**Figure 3.4** RT PCR of *rpa1*, *rpa2* and *rpa3*

RT-PCR of *rpa1* (340 bp), *rpa2* (759 bp) and *rpa3* (345 bp) using RNA isolated from wild-type *H. volcanii* H678 at different growth phases. The gene for the  $\alpha$ -subunit of the RNA polymerase *rpoA* was used as a positive control, as it is expressed constitutively. The negative control consisted of *rpoA* NO RT, where the RT step of the PCR is missed, demonstrating that the PCR product is not due to residual contaminating DNA.

**Table 3.2** *rpa1ap*, *rpe* and *rpa3ap* internal RT primers

Primer	Target gene	Product size
ARPA1P1 F	<i>rpa1ap</i>	371 bp
ARPA1P1 R	<i>rpa1ap</i>	
ARPA1P2 F	<i>rpe</i>	576 bp
ARPA1P2 R	<i>rpe</i>	
ARPA3P1 F	<i>rpa3ap</i>	330 bp
ARPA3P1 R	<i>rpa3ap</i>	



**Figure 3.5** RT PCR of *rpa1ap*, *rpe* and *rpa3ap*

RT PCR of *rpa1ap* (371 bp), *rpe* (576 bp) and *rpa3ap* (330 bp) using RNA isolated from *H. volcanii* H678 at different growth phases. The gene for the  $\alpha$ -subunit of the RNA polymerase *rpoA* was used as a positive control, as it is expressed constitutively. The negative control consisted of *rpoA* NO RT, where the RT step of the PCR is missed, demonstrating that the PCR product is transcribed from RNA not DNA.

### 3.3 Discussion

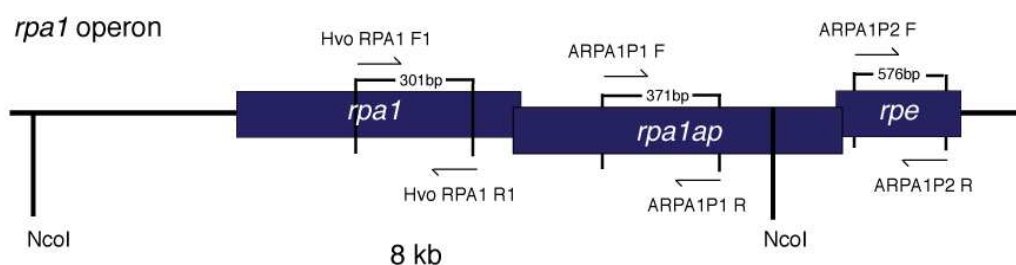
The initial findings suggest that all of the RPAs and their associated protein genes are expressed. However the relationship between the three RPAs does not seem as straightforward as seen *P. furiosus*. RPA2 did not cluster in the same clade as RPA1 and 3, signifying a level of divergence between RPA2 and RPA1/3. This is compounded by the existence of *rpa1* and *rpa3* in operons with their respective associated protein genes. Although the operon structure of RPA1 and RPA3 is not identical, their clustering in the same clade compared to that of RPA2 strongly suggests one of them may have arisen through a gene duplication event.

As discussed in Chapter 1, RPA functions in stabilising and protecting ssDNA from nuclease attack, as well as providing a platform for protein:protein interaction during DNA replication and repair. This thesis will determine if *H. volcanii* RPA1, 2 and 3 have a role in DNA replication and repair. In addition, it will be tested whether the three RPAs fulfil this role collectively as a heterotrimer, as seen in *P. furiosus* and in eukaryotes, or independently. Furthermore the role of the associated proteins will be examined; whether they have a role in DNA replication and repair, and if so does that involve acting in a complex with their respective RPAs.

## Chapter 4: RPA1

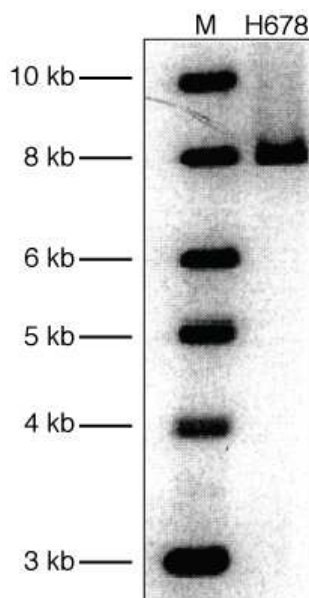
### 4.1. Isolation of the *rpa1* operon

The genomic sequence of *H. volcanii* was used to identify appropriate restriction sites to obtain a genomic clone of *rpa1*. This was carried out before it was known *rpa1* was in an operon with two other genes. Genomic DNA from *H. volcanii* strain H678 was digested originally with *Nco*I to give a predicted fragment size of 8 kb, shown in Figure 4.1, and was confirmed by Southern hybridisation shown in Figure 4.2.



**Figure 4.1.** *Nco*I genomic digest to isolate *rpa1*

The location of *Nco*I restriction sites generating a 8 kb fragment are shown, as well as the primers used to generate the 301 bp, 371 bp and 576 bp product that was used in the Southern hybridisation for *rpa1*, *rpa1ap* and *rpe* respectively. The associated gene containing an OB-fold (*rpa1ap*) and the associated phosphoesterase (*rpe*) are shown.



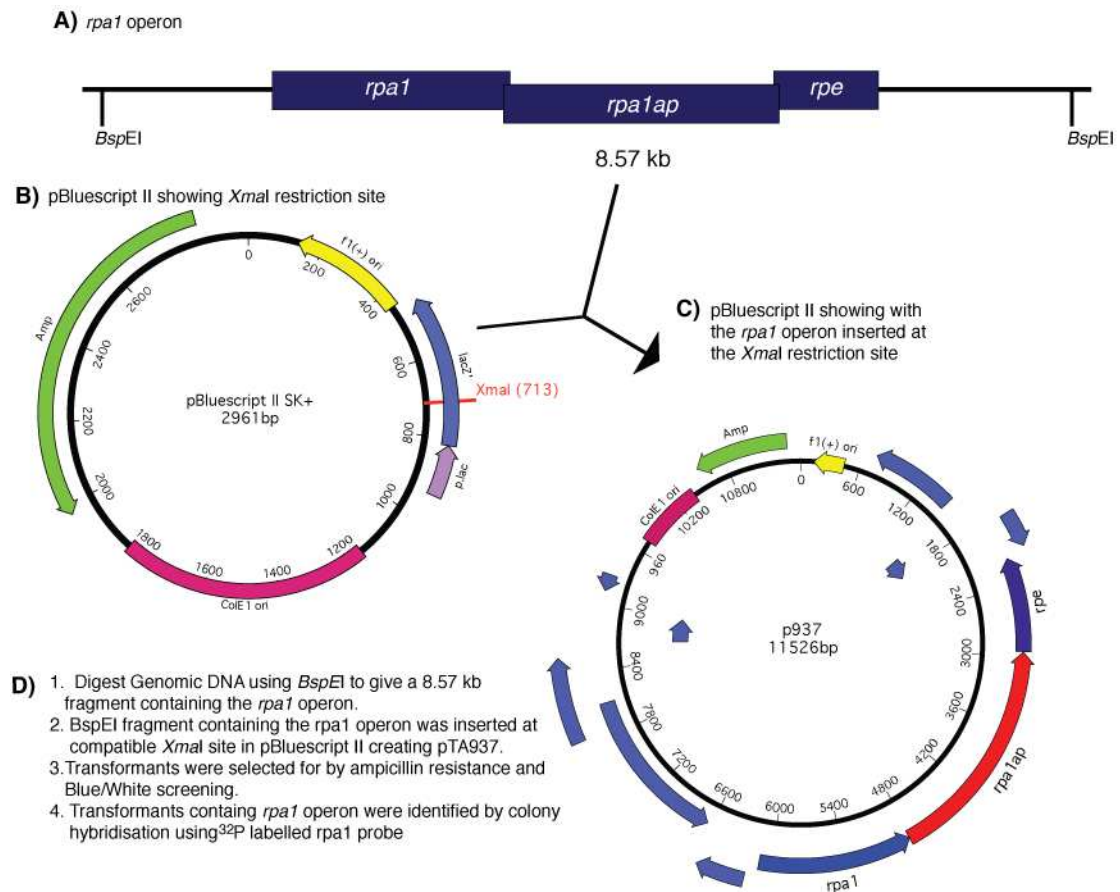
**Figure 4.2.** Southern blot analysis of *rpa1* genomic fragment

Lane M shows the size marker. Genomic DNA (H678) was digested with *Nco*I and probed with *rpa1* PCR product, shown in Figure 4.1.

However it became clear that *rpa1* is in an operon with two other genes. To isolate the whole of the *rpa1* operon, genomic DNA was digested with *Bsp*EI, giving an expected



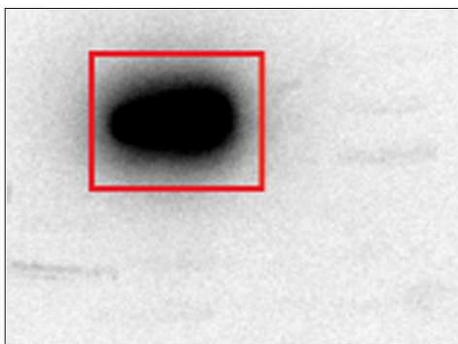
fragment size of ~8.6 kb shown in Figure 4.3. A library of *Bsp*EI DNA fragments 8-9 kb in size was cloned in the plasmid pBluescript II SK+, which features an ampicillin marker, as shown in Figure 4.3.



**Figure 4.3. Cloning strategy to isolate the *rpa1* operon**

A) *Bsp*EI restriction sites used to insert the *rpa1* operon into B) pBluescript II at the compatible restriction site, *Xma*I, to generate C) a genomic clone of the *rpa1* operon, pTA937.

The library of 8-9 kb *Bsp*EI fragments was transformed into *E. coli* XL-1, blue/white screening was used to isolate transformants with an insert. These transformants were then screened by colony hybridisation for the presence of *rpa1* using *rpa1* probe labelled with  $^{32}$ P, as shown in Figure 4.4. Plasmid extraction followed by sequencing was used to confirm the isolation of the *rpa1* operon in pTA937.



**Figure 4.4. Colony hybridisation screening for *rpa1* operon.**

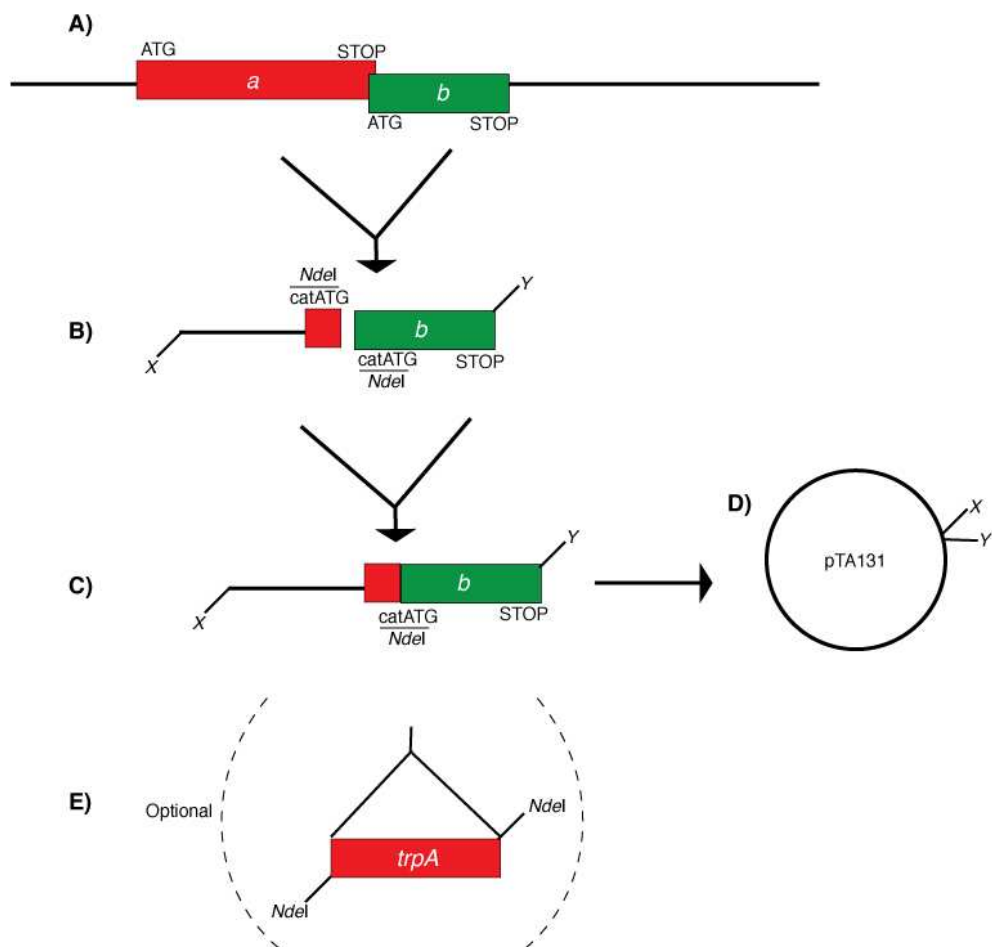
Screening of transformants containing *rpa1* operon by colony hybridisation with  $^{32}\text{P}$  labelled *rpa1* probe (See Figure 4.1).

## 4.2. Generation of single *rpa1*, *rpa1ap* and *rpe* deletion mutants

The generation of deletion mutants allows the study of a DNA replication or repair phenotype. Upstream and downstream region of the gene to be deleted were amplified, novel restriction sites were introduced to allow ligation of the two regions into the plasmid pTA131, which features a *pyrE2* and an ampicillin marker. However due to the structure of the *rpa1* operon, the restriction sites must incorporate the start codon of the gene to be deleted in order to maintain the reading frame of *rpa1ap* and *rpe*, and their start and stop codons, thus avoiding a polar effect on downstream *rpa1ap*, *rpe* genes shown in Figure 4.5. For example the deletion construct of *rpa1* will contain an internal *NdeI* restriction site that cuts at the *rpa1* start codon, and the downstream region will contain the start codon of *rpa1ap* followed by the stop codon of *rpa1*, as well as a *NdeI* site. This generates a ‘clean’ deletion so not as to interfere with the expression of the downstream *rpa1ap* and *rpe* genes.

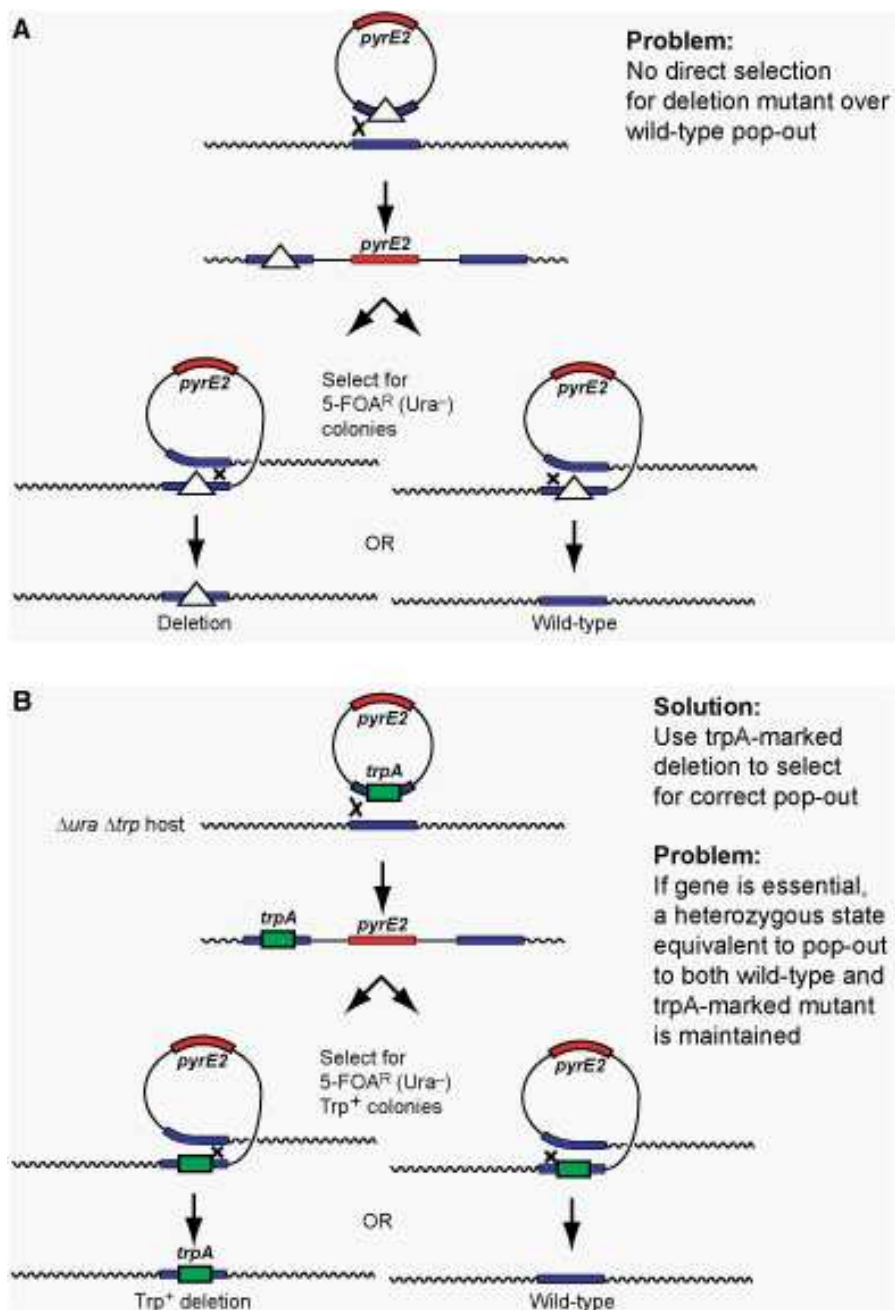
The plasmid was then transformed into *E. coli* XL-1, plasmid extraction of transformants followed by restriction digests was used to screen transformants. The deletion construct was confirmed by sequencing and transformed in *E. coli dam*- before being transformed into the *H. volcanii* strain H195, selecting for uracil phototrophy (*pyrE2*<sup>+</sup>). Transformation into *E. coli dam* mutant is required due to the presence of restriction enzymes in *H. volcanii* that target 5'-G<sup>me</sup>ATC-3' sites generated by the Dam methylase (Hartman, Norais *et al.* 2010). By passing plasmid constructs through *E. coli dam* mutants, 5'-G<sup>me</sup>ATC-3' sites are converted to 5'-GATC-3' (Leigh, Albers *et al.* 2011). The plasmid containing the deletion construct integrated onto the chromosome through recombination between upstream and downstream regions of the plasmid and homologous chromosomal regions, generating a pop-in shown in Figure 4.6. The pop-in strain was grown non-selectively in Hv-YPC

to allow recombination between homologous regions of plasmid and chromosome, resulting in the pop-out of the plasmid from the chromosome. This was selected for by 5-FOA resistance. Depending on the orientation of the second recombination event relative to the first, the resulting strain is either deleted for desired gene or wild-type. The insertion of an selectable marker, such as *trpA*, between the upstream and downstream region, can allow bias for the generation of a deletion over wild-type using selection for tryptophan phototrophy. This method of generating a genomic deletion in *H. volcanii* is known as the counter-selective pop-in/pop-out method, shown in Figure 4.6. (Allers, Ngo *et al.* 2004).



**Figure 4.5. Generic method for deleting a gene in an operon**

A) Operon containing genes *a* and *b*, the former of which is to be deleted. B) Internal *NdeI* sites are introduced at the start codons (ATG) of each gene and external restriction sites (X and Y). C) *NdeI* sites are used to ligate the upstream and downstream regions. D) The ligated product is inserted into pTA131 using external restriction sites. E) Optionally the selectable marker *trpA* can be inserted using internal *NdeI* site, amplified by PCR from pTA298, see (Figure 4.13).



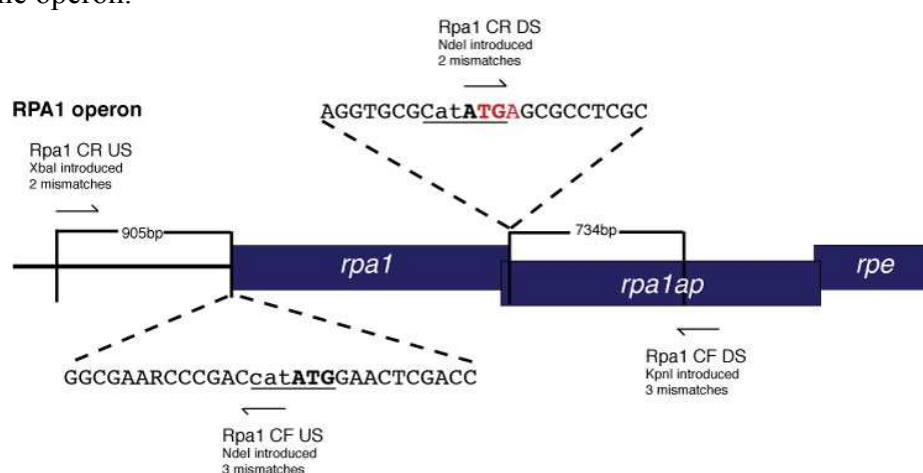
**Figure 4.6. Generating a gene knockout**

A) Generating a deletion mutant without a selectable marker. B) Generating a mutant with a selectable marker. See text for details.

### 4.2.1. Construction of *rpa1* deletion mutant

#### Design of primers

Upstream and downstream regions of *rpa1* were amplified from the genomic *rpa1* clone pTA937 by touchdown PCR using the primers shown in Table 4.1, introducing external *KpnI* and *XbaI* restriction sites and an internal *NdeI* restriction site, shown in Figure 4.7. The restriction site *NdeI* was chosen, since it contains ATG, to avoid changing the nucleotide A after the start codon (ATG) and altering the reading frame of the operon.



**Figure 4.7 Introduction of internal *NdeI* sites**

The introduction of internal *NdeI* restriction sites whilst maintain the reading frame and thus the start codon of the downstream *rpa1ap*. Start codons shown in bold.

Primer	Sequence (5'-3')	Site inserted
Rpa1 CF DS	GTTTCGAGGT <u>acc</u> GTTTCGGGGAGC	<i>KpnI</i>
Rpa1 CR DS	AGGTGCGC <u>at</u> <b>ATG</b> AGCGCCTCGC	<i>NdeI</i>
Rpa1 CR US	TACTACGTCTaGaCGGACCTGTTTCG	<i>XbaI</i>
Rpa1 CF US	GGTCGAGTTC <b>CAT</b> atgGTCGGGATTTCGCC	<i>NdeI</i>

**Table 4.1 Design of deletion primers for *rpa1***

The sequence for each deletion primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of the upstream (US) and downstream (DS) regions. Start codon of the *rpa1ap* in the Rpa1 CR DS primer is shown in bold as is the start codon for *rpa1* in the Rpa1 CF US primer. The stop codon of *rpa1* is shown in blue in the Rpa1 CR DS primer.

### **Construction of *rpa1* deletion plasmid**

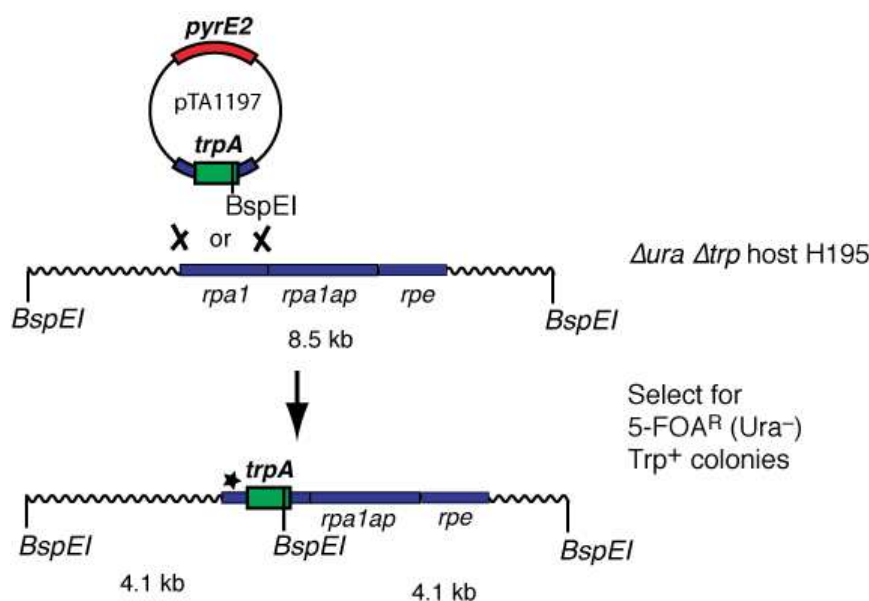
The PCR products of the upstream and downstream regions were gel purified, digested with *NdeI* and ligated together. The upstream/downstream ligation was ligated into the plasmid pTA131 using the *KpnI* and *XbaI* sites, and was transformed into *E. coli* XL-1 strain, as shown in Figure 4.8. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1141 construct.





### Pop-in/ pop-out of the deletion construct

Transformation into the *E. coli dam* mutant generated the plasmid pTA1197, which was transformed into *H. volcanii* H1201 (pop-in) to generate a genomic deletion of *rpa1*, H1217, as shown in Figure 4.9.

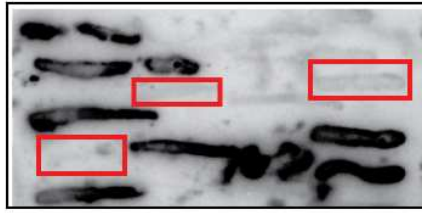


**Figure 4.9. Schematic diagram of *rpa1* deletion construction**

The deletion plasmid pTA1197 was constructed as described in Figure 4.8. Restriction digest with *BspEI* and the resulting fragments expected that were used in the Southern blot analysis are shown. Integration onto the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) whilst selecting for *trpA*<sup>+</sup> resulted in the  $\Delta rpa1$  strain. The region that the <sup>32</sup>P probe hybridises to is denoted by a star.

Due to the polyploidy of *H. volcanii* it is possible for chromosome copies containing wild-type *rpa1* and chromosome copies containing the *rpa1* deletion to co-exist in one strain, termed a merodiploid strain. This is more likely to occur if the deletion construct is marked with a selectable marker, preventing all of the chromosome copies reverting back to the wild-type gene, in this case *rpa1*. In order to avoid merodiploid *H. volcanii* strains, potential deletions were screened by colony hybridisation, using a radioactive <sup>32</sup>P labelled *rpa1* probe (see Figure 4.1), pop-outs that did not hybridise at all with the probe were selected as potential deletion mutants that were not merodiploid, shown in Figure 4.10.

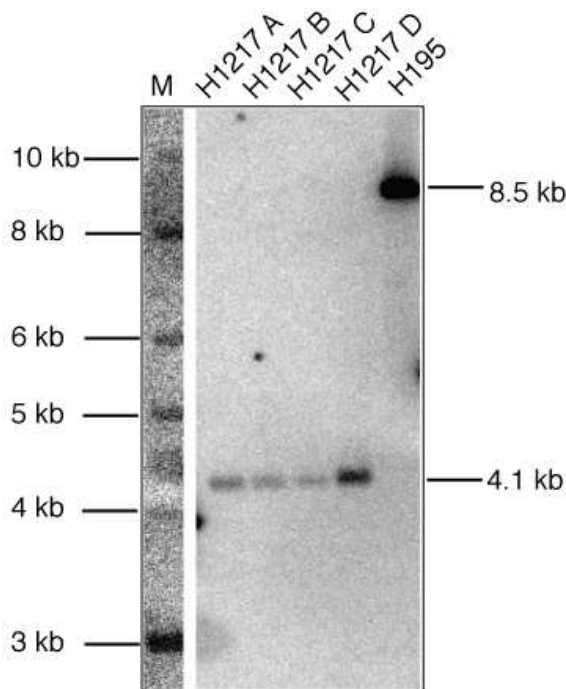




**Figure 4.10 Colony lift and hybridisation of H1201 popouts**

Pop-outs (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv-Ca +thy +ura, to maintain selection for *trpA* marked deletion construct in case of merodiploidy. Three popout candidates that failed to hybridise with the *rpa1* probe are indicated by red boxes.

The genomic sequence of *H. volcanii* was used to select appropriate restriction enzymes to digest the genomic DNA of the potential deletion mutants, which was then used in Southern hybridisation to confirm genomic deletion of *rpa1*, shown in Figure 4.11. A <sup>32</sup>P labelled probe of the upstream, autotrophic marker and downstream region was used in the hybridisation, shown in Figure 4.8.



**Figure 4.11 Southern blot analysis of H1201 popouts (H1217)**

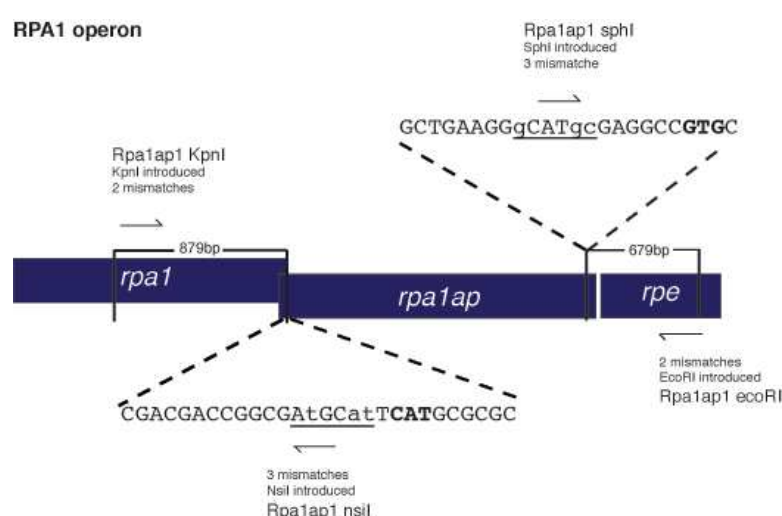
Lane M shows size marker lane. Genomic DNA was digested with *BspE1* and probed with the flanking regions and the *trpA* marker of the deletion construct p1141 *XbaI/KpnI*. Genomic DNA of H195 was used a wild-type control.

#### **4.2.2 Construction of *rpa1ap* deletion mutant**

##### **Design of primers**

Upstream and downstream regions of *rpa1ap* were amplified from the genomic *rpa1* clone pTA937 by touch down PCR using the primers shown in Table 4.2, introducing

external *KpnI* and *EcoRI* and internal *NsiI* and *SphI* restriction sites. In order to maintain the *rpaI* sequence, whose stop codon occurs after the start codon of *rpaIap*, the internal restriction site cannot be *NdeI* as this would require a base change in the *rpaI* sequence. To overcome this a three step procedure utilized the *trpA* marker from pTA298, which was inserted using *PstI* and *SphI* sites. The presence of the *trpA* marker in the *rpaIap* deletion construct will allow the biasing of generating a deletion mutant over a wild-type as shown in Figure 4.15, as well as allowing the use of Hv-Ca as discussed previously.



**Figure 4.12 Introduction of internal *NdeI* sites**

The introduction of internal *NdeI* restriction sites whilst maintain the reading frame and thus the start codon of the downstream *rpe*. Start codons are shown in bold

Primer	Sequence	Site inserted
Rpa1ap1 sphI ds	GCTGAAGGGC <b>atTGC</b> <u>G</u> AGGCCGTGC	<i>SphI</i>
Rpa1ap1 ecoRI ds	CGGCGAGAGAA <b>tt</b> <u>CCCTG</u> CCCCGGG	<i>EcoRI</i>
Rpa1ap1 kpnI us	CCGCGAGTGG <b>t</b> <u>ACC</u> GCAAGCCCG	<i>KpnI</i>
Rpa1ap1 nsiI us	CGACGACCGGCGA <b>t</b> <u>GCat</u> <b>TCAT</b> GCGCGC	<i>NsiI</i>

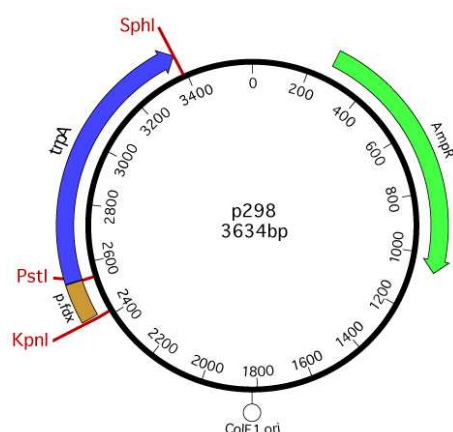
**Table 4.2 Primers for the construction the *rpaIap* deletion construct**

Primers used to generate the *rpaIap* deletion construct are shown. The sequence for each deletion primer is shown with the base changes (in lower case) introducing a novel restriction site (underlined) to allow cloning of the upstream (us) and downstream (ds) regions. The stop codon for *rpaI* is shown in blue and the start

codon for *rpaIap* is shown in bold in the RpaIap1 *nsiI* us primer. Part of the stop codon for *rpaIap* is shown in the primer RpaIap1 *sphI* ds in blue.

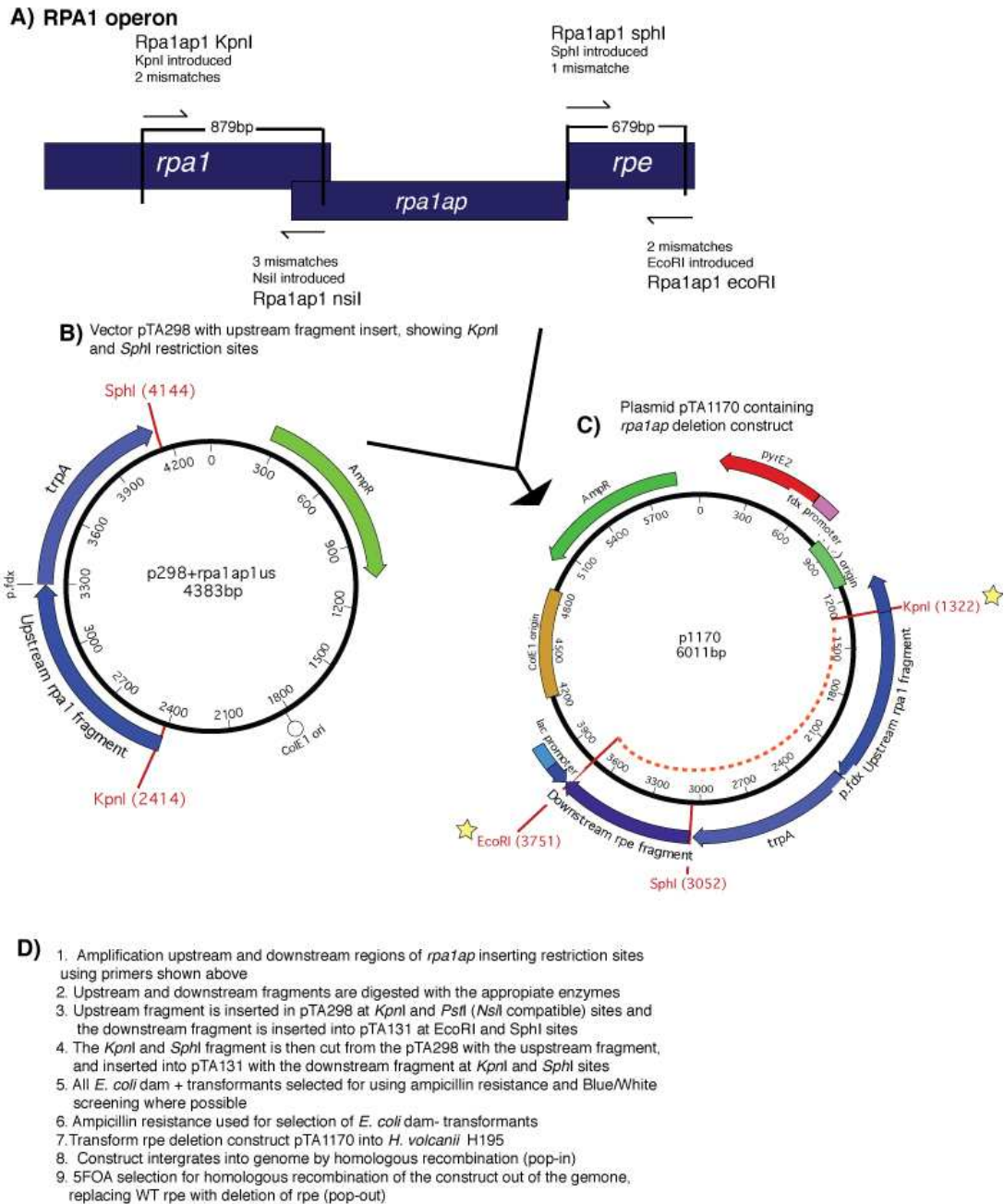
### Construction of *rpaIap* deletion plasmid

The upstream flanking region of *rpaIap* was amplified introducing an external *KpnI* site and an internal *NsiI*, which maintains the *rpaI* coding sequence and is compatible with the *PstI* site in pTA298. This upstream region was gel purified, digested with the appropriate enzymes and inserted into the digested and gel purified pTA298, shown in Figure 4.13. A downstream region was amplified inserting an internal *SphI* site and an external *EcoRI* site. The upstream region and the *trpA* marker fragment were then digested from the plasmid with *SphI* and ligated with the downstream region. This ligation, was inserted into pTA131 at *KpnI* and *EcoRI* restriction sites. Due to the sequence constraints the first base of the *rpaIap* stop codon had to be changed from a T to a C, however this does not affect the sequence of downstream *rpe* gene. The plasmid pTA1170 was then transformed into *E. coli* XL-1 strain. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct construct pTA1170, as shown in Figure 4.14.



**Figure 4.13 pTA298**

Plasmid map of pTA298 showing *PstI*, *SphI* and *KpnI* restriction sites used in Figure 4.14. Also shows *trpA* marker used in Figure 4.8.

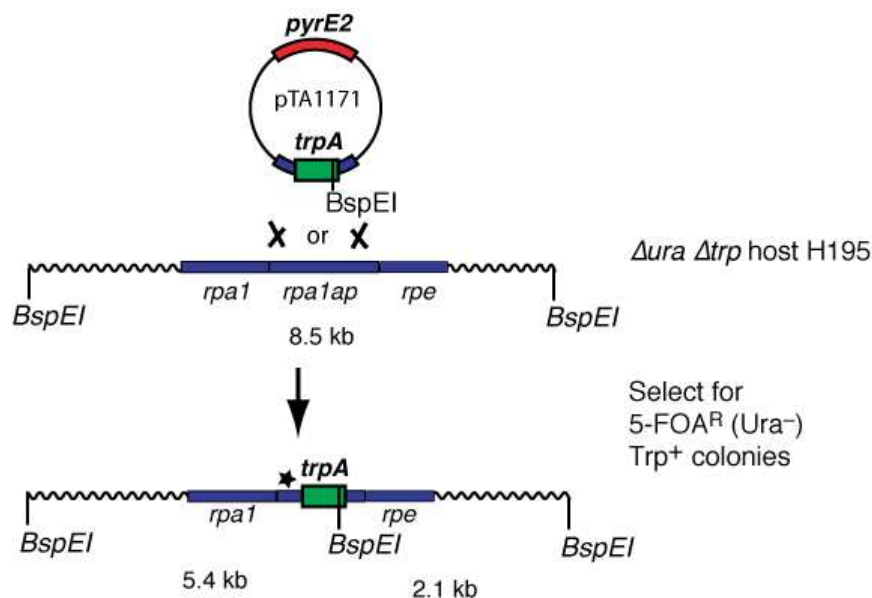


**Figure 4.14 construction of the *rpa1ap* deletion**

Figure shows the cloning technique used to generate the deletion construct of *rpa1ap*, the binding sites of the primers are shown A) the restriction site in the vector pTA298 used to insert the upstream fragment B) and the restriction sites used to insert the upstream *trpA* marker and the downstream fragment resulting in the deletion construct C). A basic sequence of construction is shown in D). Orange dotted line and yellow stars denotes region that was used as a <sup>32</sup>P labelled probe in Southern hybridisation shown in Figure 4.17.

### Pop-in/ pop-out of the deletion construct

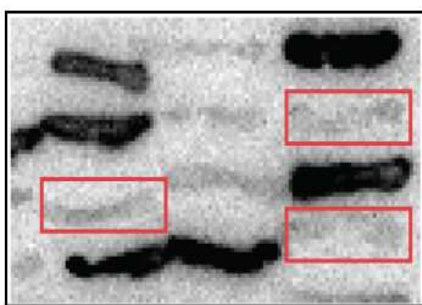
Following transformation into the *E. coli dam*<sup>-</sup> strain the plasmid pTA1170 was transformed into *H. volcanii*, generating H1198 (pop-in) to generate a genomic deletion of *rpa1ap*, as shown in Figure 4.15.



**Figure 4.15 Schematic diagram of *rpa1ap* deletion construction**

The deletion plasmid pTA1171 was constructed as described in Figure 4.14. Restriction digest with *BspEI* (see text) and the resulting fragments expected are shown. Integration onto the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) whilst selecting for *trpA*<sup>+</sup> resulted in the  $\Delta rpa1ap$  strain. The region that the <sup>32</sup>P labelled probe hybridises to is denoted by a star.

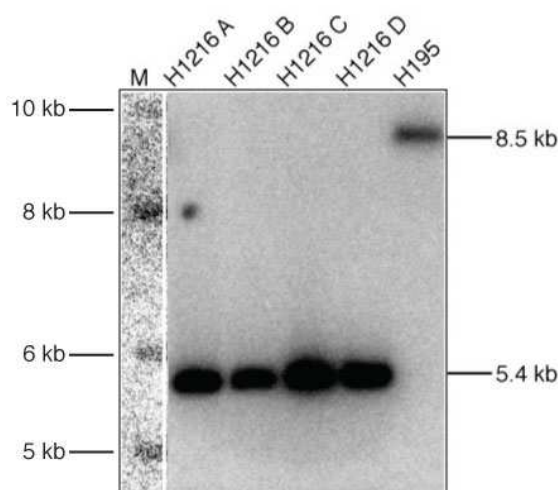
Potential deletions were screened by colony hybridisation, using a <sup>32</sup>P labelled *rpa1ap* probe, pop-outs that did not hybridise to the probe were selected as potential deletion mutants, as shown in Figure 4.16.



**Figure 4.16 Colony lift and hybridisation of H1198 popouts**

Popouts (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv-Ca +thy +ura, to maintain selection for *trpA* marked deletion construct. Red boxes indicate three popout candidates that failed to hybridise to the *rpa1ap* probe.

The genome sequence of *H. volcanii* was used to determine an appropriate restriction digest of the potential deletion mutants, this was then used in Southern hybridisation to confirm genomic deletion of *rpalap*. A  $^{32}\text{P}$  labelled probe of the upstream, *trpA* marker and downstream region was used in the hybridisation, as shown in Figure 4.17.



**Figure 4.17 Southern blot analysis of H1201 popouts (H1217)**

Lane M shows size marker. Genomic DNA was digested with *BspE1* and probed with the flanking regions and the *trpA* marker of the deletion construct p1170 *EcoRI/KpnI*. Genomic DNA of H195 was used a wild-type control. For digest see Figure 4.15.

#### 4.2.3. Construction of *rpe* deletion mutant

##### Design of primers

Upstream and downstream regions of *rpe* were amplified from the genomic clone by touchdown PCR using the primers shown in Table 4.3, introducing external *KpnI* and *EcoRI* restriction sites and an internal *BamHI* restriction site.

Primer	Sequence	Site inserted
PEbamHI F (DS)	CGACGAG <u>G</u> AtCcACCGTTTCG	<i>BamHI</i>
PEkpnI R (US)	GCTCG <u>gt</u> ACCGAGTCGAGCCC	<i>KpnI</i>
PEecorI F (DS)	GCCCCAaTTCCGTCTGATTG	<i>EcoRI</i>
PEbamHI R (US)	CTCTCGCgGaTCCTGCCCCG	<i>BamHI</i>

**Table 4.3 Primers for the construction the *rpe* deletion construct**

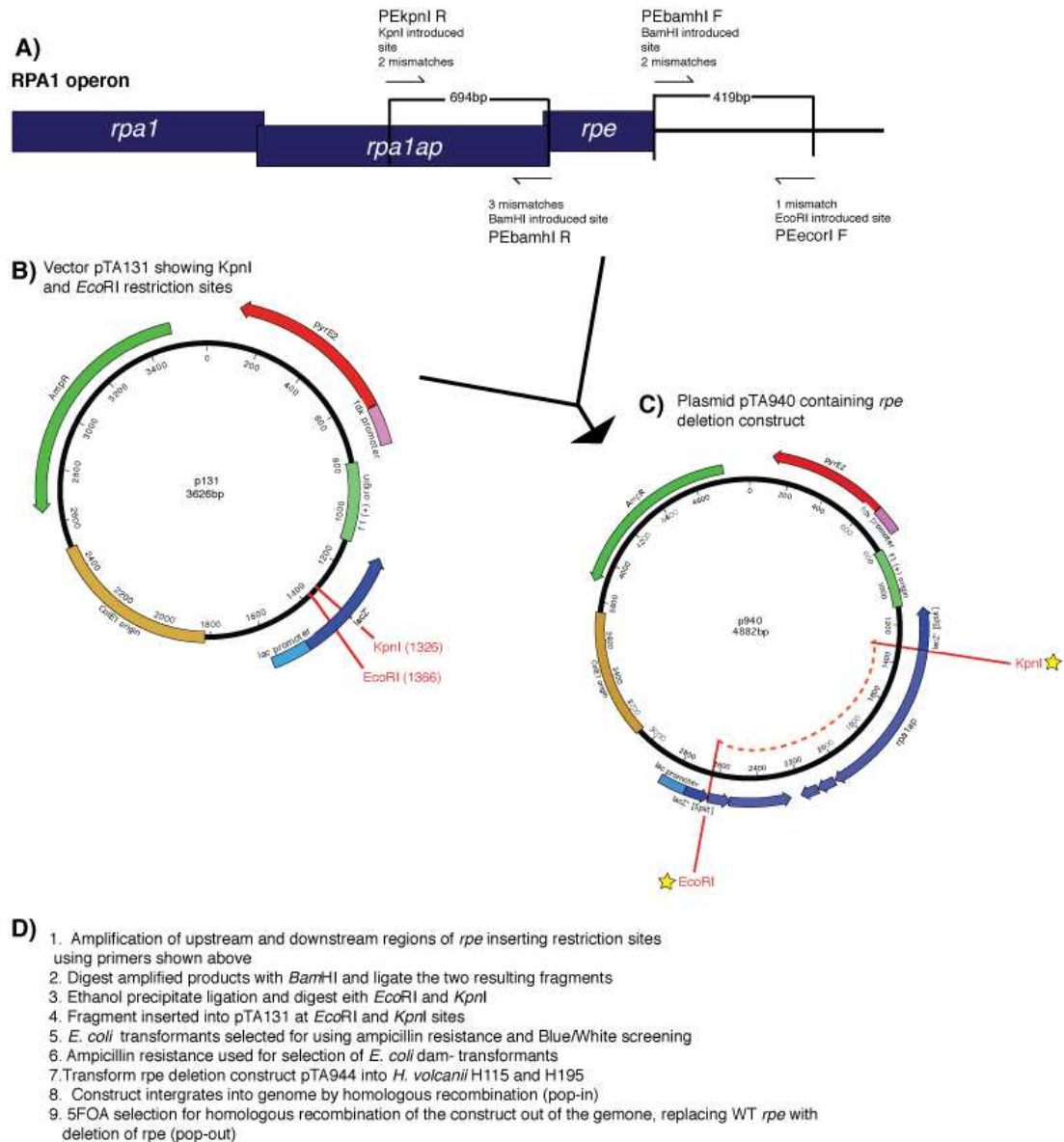
Primers used to generate the *rpe* deletion construct are shown. The sequence for each deletion primer is shown with the base changes (lower case) made to introduce a novel restriction site (underlined). PEbamHI F and PEkpnI R primers were used to clone the upstream fragment and downstream region is cloned using the primers PEbamHI R and PEecorI.

### **Construction of the *rpe* deletion plasmid**

The PCR products of the upstream and downstream regions were gel purified, digested with the appropriate restriction enzymes and ligated together using the *Bam*HI sites. The upstream/downstream ligation was then ligated into the plasmid pTA131 using the *Eco*RI and *Kpn*I sites, which was transformed into *E. coli* XL-1, as shown in Figure 4.18. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct construct, pTA940.

Following transformation into the *E. coli* the plasmid pTA940 was transformed into *H. volcanii* generating the pop-in strain H936, which was then used to generate a genomic deletion of *rpe* H1134, as shown in Figure 4.19.

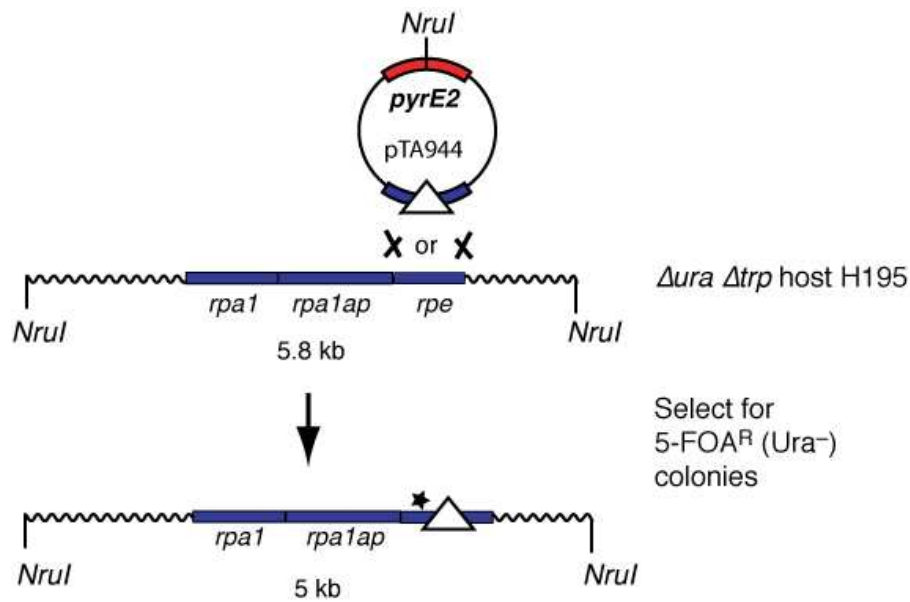




**Figure 4.18 Construction of the *rpe* deletion**

Figure shows the cloning technique used to construct the deletion construct of *rpe*, the binding sites of the primers are shown A) the restriction site in the vector used to introduce the upstream and downstream ligation B) and the resulting deletion construct pTA940 C). A basic sequence of construction is shown in D) Orange dotted line and yellow stars denotes region that was used as a <sup>32</sup>P labelled probe in Southern hybridisation shown in Figure 4.21.

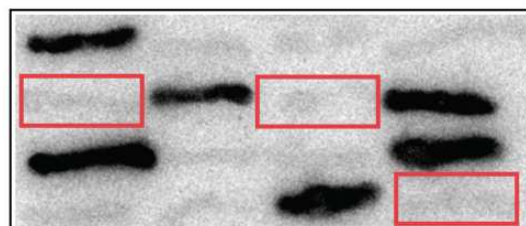




**Figure 4.19 Schematic diagram of *rpe* deletion construction**

The deletion plasmid pTA944 was constructed as described in Figure 4.18. Restriction digest with *Nru*I (see text for details) and the resulting fragments expected are shown. Integration onto the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) resulted in the  $\Delta rpe$  strain. The region that the  $^{32}\text{P}$  labelled probe hybridises to is denoted by a star.

Potential deletions were screened by colony hybridisation, using a  $^{32}\text{P}$  labelled *rpe* probe (see Figure 4.1), pop-outs that did not hybridise to the probe were selected as potential deletion mutants.

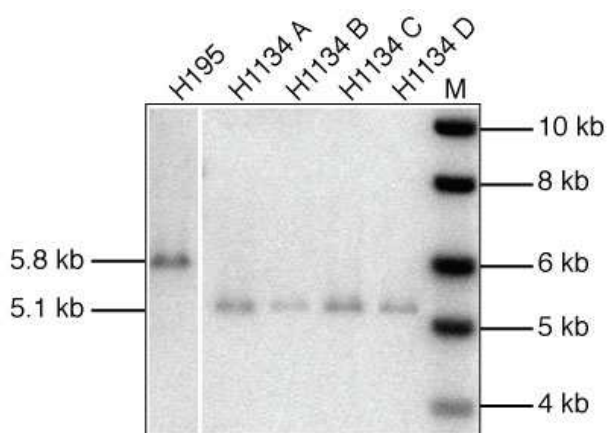


**Figure 4.20 Colony lift and hybridisation of H936 popouts**

Popouts (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv- YPC. Three popout candidates that failed to hybridise with the *rpe* probe are indicated by red boxes.

The genome sequence of *H. volcanii* was used to determine an appropriate restriction digest to screen potential deletion mutants, this was then used in a Southern hybridisation to confirm genomic deletion of *rpe*, shown in Figure 4.21. A  $^{32}\text{P}$  labelled

probe of the upstream, *trpA* and downstream region of *rpaI* was used in the hybridisation, as shown in Figure 4.18.



**Figure 4.21 Southern blot analysis of H936 popouts (H1134)**

Lane M shows the size marker. Genomic DNA was digested with *NruI* and probed with the flanking regions of the deletion construct p940 *EcoRI/KpnI*. Genomic DNA of H195 was used as a wild-type control. See Figure 4.19 for further digest details.

### 4.3. Generation of *rpa1* operon deletion

#### 4.3.1 Construction of *rpa1* operon deletion mutant

##### Design of primers

Upstream and downstream regions of *rpa1* operon were amplified from the genomic *rpa1* clone pTA937 by touchdown PCR using the primers shown in Table 4.4, introducing external *Eco*RI and *Xba*I and an internal *Nde*I restriction sites. Primers RPA1 CR US and PEcoRI were utilized from the previous generation of *rpa1* and *rpe* deletion mutants respectively.

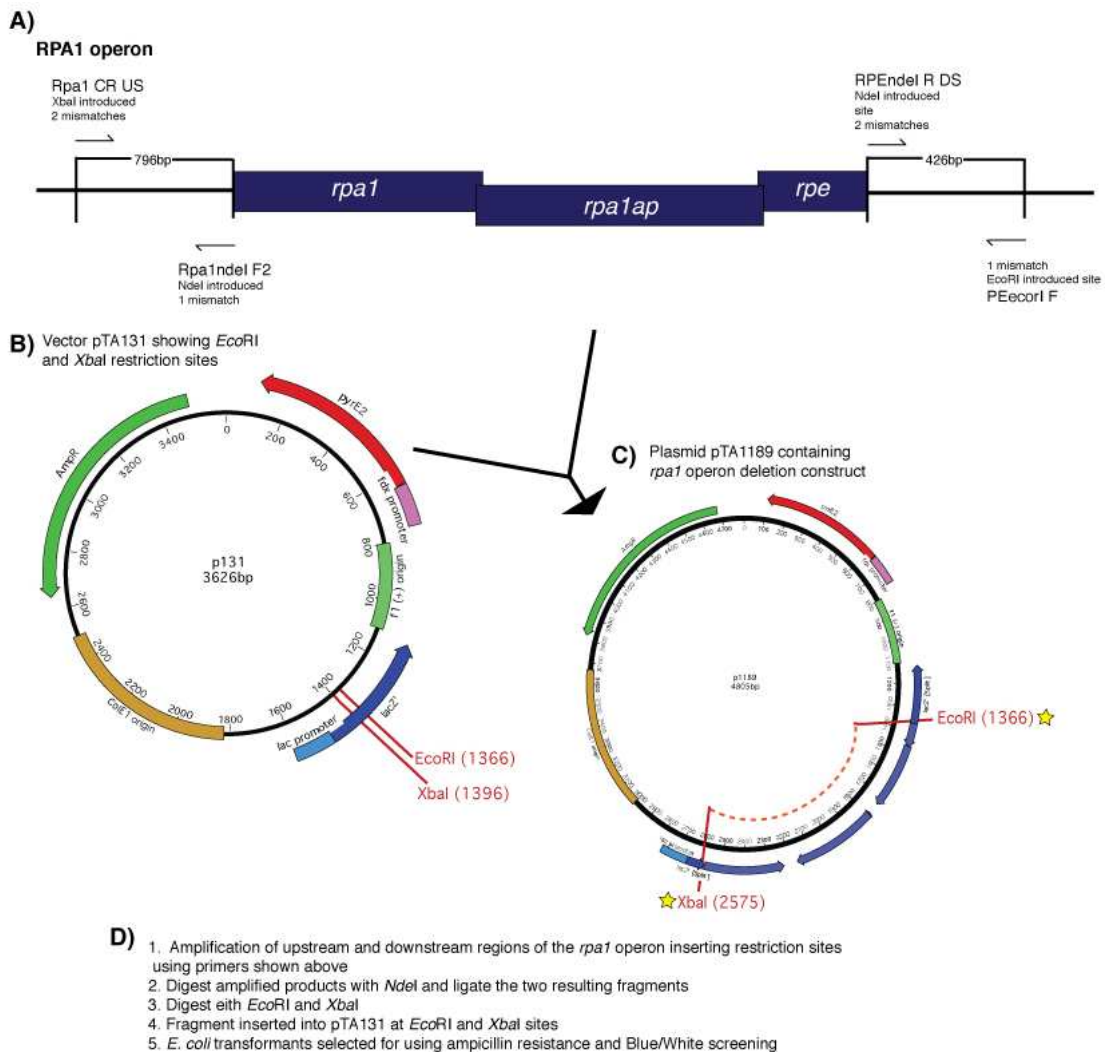
Primer	Sequence	Restriction site inserted
Rpa1 CR US	TACTACGTCTaGaCGGACCTGTTCG	<i>Xba</i> I
Rpa1NdeI F2	GTTGGACCCaTATGTCGACGACG	<i>Nde</i> I
PEcoRI	GCCCGAaTTCCGTCTGATTG	<i>Eco</i> RI
RPEndeI R DS	CTACCGGAACaTa <b>TG</b> ACTCGGGTCG	<i>Nde</i> I

**Table 4.4. Primers for the construction the *rpe* deletion construct**

Primers used to generate the *rpa1* operon deletion construct are shown. The sequence for each deletion primer is shown with the base changes (lower case) made to introduce a novel restriction site (underlined). Primers Rpa1 CR US and Rpa1NdeI F2 were used to clone the upstream fragment and downstream region was cloned using the primers PEcoRI and RPEndeI R DS. The stop codon of the *rpe* is shown in bold.

##### Construction of the *rpa1* operon deletion plasmid

The PCR products of the upstream and downstream regions were gel purified, digested with the appropriate restriction enzymes and ligated together using the *Nde*I sites. The upstream/downstream ligation was then ligated into the plasmid pTA131 using the *Eco*RI and *Xba*I sites, which was transformed into *E. coli* XL-1, as shown in Figure 4.22. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct construct.

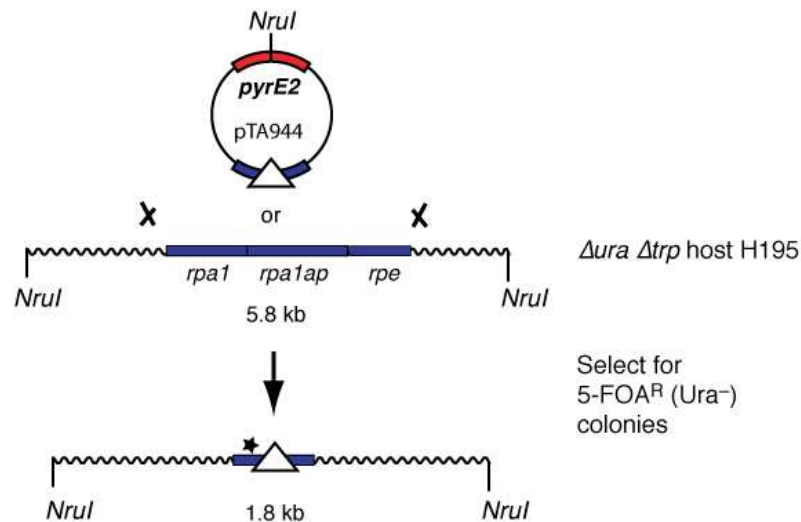


**Figure 4.22 Construction of the *rpa1* operon deletion**

Figure shows the cloning technique used to generate the deletion construct of *rpa1* operon, the binding sites of the primers are shown A) the restriction site in the vector used to insert the upstream and downstream ligation B) and the resulting deletion construct pTA1181 C). A basic sequence of construction is shown in D) Orange dotted line and yellow stars denotes region that was used as a  $^{32}\text{P}$  labelled probe in Southern hybridisation shown in Figure 4.25.

### Pop-in/ pop-out of the deletion construct

Following transformation into the *E. coli dam*- the plasmid pTA1191 was transformed into H195 generating the pop-in H1245, which was used to generate a genomic deletion of *rpa1* operon (H1246), as shown in Figure 4.23.



**Figure 4.23. Schematic diagram of *rpe* deletion construction**

The deletion plasmid pTA1191 was constructed as described in Figure 4.22. Restriction digest with *NruI* (see text for details) and the resulting fragments expected are shown. Integration onto the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) resulted in the  $\Delta rpa1$  operon strain. The region that the  $^{32}\text{P}$  labelled probe hybridises to is denoted by a star.

Potential deletions were screened by colony hybridisation, using  $^{32}\text{P}$  labelled *rpa1* probe (see Figure 4.1), pop-outs that did not hybridise to the probe were selected as potential deletion mutants, as shown in Figure 4.24.

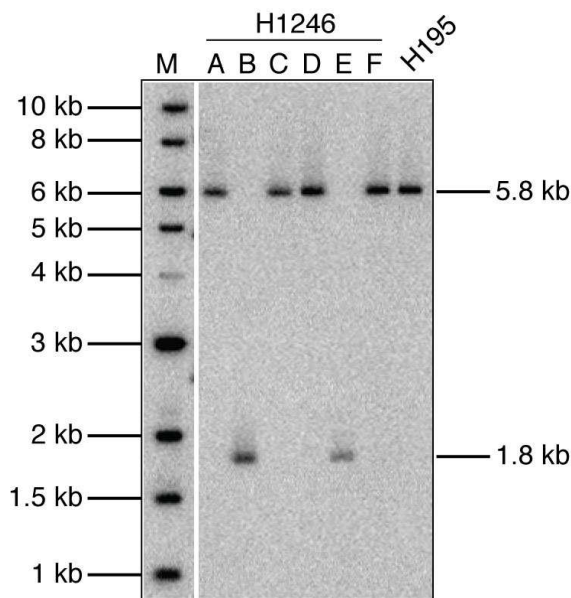


**Figure 4.24 Colony lift and hybridisation of H1245 popouts**

Popouts (*ura*-) selected for on 5FOA were patched out on Hv- YPC. Four popout that failed to hybridise with *rpa1* probe candidates are indicated by red boxes.

The genomic sequence of *H. volcanii* was used to determine an appropriate restriction digest to screen the potential deletion mutants, this was then used in a Southern hybridisation to confirm genomic deletion of *rpa1* operon. A  $^{32}\text{P}$  labelled probe of the

upstream, *trpA* and downstream region of the *rpa1* operon (see Figure 4.22) was used in the hybridisation, shown in Figure 4.25.



**Figure 4.25 Southern blot analysis of H1245 popouts (H1246)**

Lane M shows size marker lane. Genomic DNA was digested with *NruI* and probed with the flanking regions of the deletion construct pTA1206 digested with *EcoRI/XbaI*. Genomic DNA of H195 was used a wild-type control. See Figure 4.23 for further restriction digest details.

#### 4.4. DNA damage assays of deletion mutants

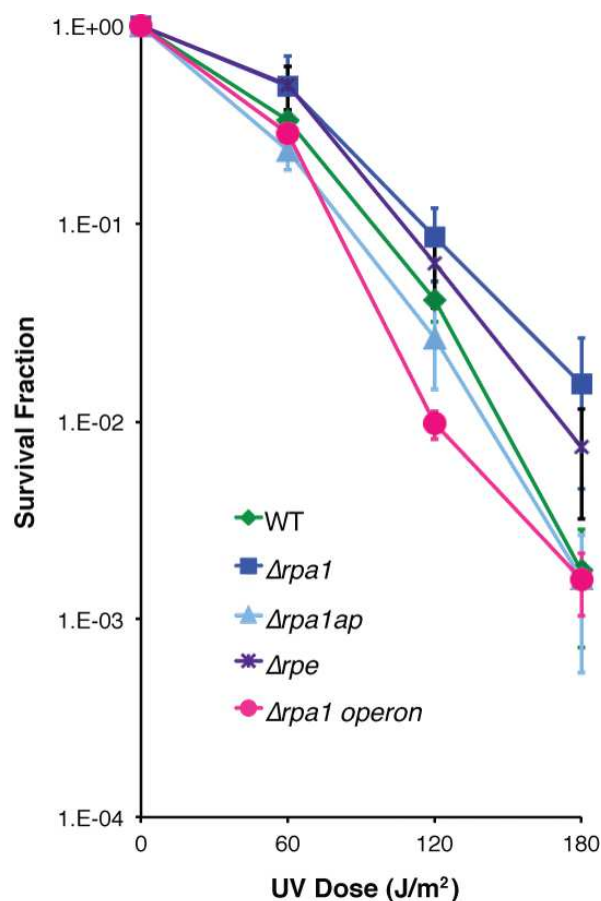
To examine if RPA1, RPA1AP or RPE play a role in DNA repair, as is true for both the bacterial SSB and eukaryotic RPA, DNA damage assays were performed using the single deletion mutants. UV irradiation was used to examine the effects of the formation of 6-4 pyrimidine-pyrimidone dimer photoproducts, which result indirectly in DSBs, requiring homologous recombination or single-strand annealing DNA damage repair pathways (Fousteri and Mullenders 2008) (Rouillon and White 2011). Mitomycin C (MMC), a chemotherapeutic DNA cross-linking agent (Tomasz, Lipman *et al.* 1987), was also used to study the effect of cross-links and their removal by nucleotide excision repair (NER) and homologous recombination (HR). In both cases, if the damaged DNA were to be left unrepaired then this would result in stalled replication forks, which if the cell was unable to restart after repair would cause cell death. Hydrogen peroxide ( $H_2O_2$ ) causes oxidative damage that can result in ssDNA breaks (Freidberg, Walker *et al.* 2006). In eukaryotes, both pathways require RPA action to bind and stabilise the ssDNA, allowing the progression of repair. It is possible that the RPE is required to act directly on damaged DNA, using its phosphoesterase activity to resect ‘dirty’ DNA ends. RPE features a calcineurin-like phosphoesterase domain, which could provide RPE with a nuclease activity that could

function in the processing of DNA ends at a DNA break. This is seen with the Mre11 in eukaryotes, which possesses 3'-5' exonuclease and endonuclease activity and makes the initial incisions in the processing of the DNA ends at a DSB (Symington 2002).

#### **4.4.1. UV DNA damage assays**

UV irradiation of cells induces the formation of cyclobutane-pyrimidine dimers (CPDs) and pyrimidine-pyrimidone photoproducts (6-4PPs), both of which result in the distortion of the DNA helix and stall DNA replication forks. These DNA lesions can be repaired by direct photoreactivation and NER in bacteria and most eukaryotes, and NER alone in the placental mammals. Defects in human NER result in xeroderma pigmentosum (XP), Cockayne's syndrome (CS) and trichothiodystrophy (TTD) (Rouillon and White 2011) (Fousteri and Mullenders 2008). *H. volcanii* possesses both eukaryotic (*rad1*/XPF, *rad2*/XPG, *rad3*/XPD and *rad25*/XPB) and bacterial (*uvrA*, *uvrB*, *uvrC* and *uvrD*) NER homologs (Lestini, Duan *et al.* 2010) (Duan 2008).

To examine if RPA1, RPA1AP and RPE function in these pathways, particularly HR, single deletion mutants and the *rpa1* operon deletion mutant were analysed in parallel with H195 (*rpa1*<sup>+</sup>, *rpa1ap*<sup>+</sup>, *rpe*<sup>+</sup>) in a UV assay, shown in Figure 4.26. The data show none of the single deletion mutants or the *rpa1* operon mutant are significantly sensitive to UV, indicating that neither RPA1, RPA1AP nor RPE are involved significantly in the removal and repair of CPDs and 6-4PPs.



**Figure 4.26 UV sensitivity of *rpa1* deletion mutants**

The survival fraction of wild-type (WT) H195, H1217 ( $\Delta rpa1$ ), H1216 ( $\Delta rpa1ap$ ), H1134 ( $\Delta rpe$ ) and H1246 ( $\Delta rpa1$  operon). The data shown is mean and standard error of three repeats.

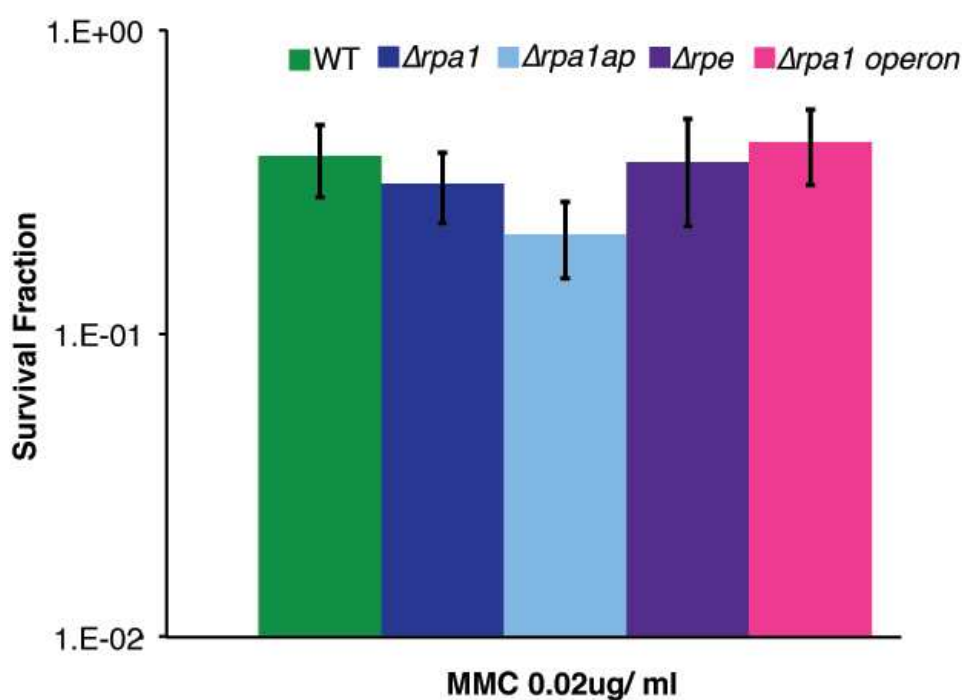
#### **4.4.2. MMC DNA damage assays**

Mitomycin C (MMC) is a bifunctional alkylating and chemotherapeutic agent that reacts covalently with DNA, forming interstrand crosslinks (ICLs) with guanine residues of 5'-CG-3' sequences. The crosslinks prevent the duplex DNA from being unwound, thus inhibiting DNA processes including DNA replication and recombination. As discussed earlier, ICLs can be removed and the DNA repaired by NER and HR pathways, which involves generating a dsDNA break intermediate (Tomasz, Lipman *et al.* 1987).

To examine if RPA1, RPA1AP and RPE function in these pathways, particularly HR, single deletion mutants and the *rpa1* operon deletion mutant were analysed in MMC



assay, shown in Figure 4.27. Hv-YPC +thy plates containing 0.02 µg/ml MMC were used in parallel with H195 (*rpa1*<sup>+</sup>, *rpa1ap*<sup>+</sup>, *rpe*<sup>+</sup>). The data shows none of the single deletion mutants or the *rpa1* operon mutant are significantly sensitive to MMC, indicating that neither RPA1, RPA1AP nor RPE are involved significantly in the removal and repair of ICLs.



**Figure 4.27 MMC sensitivity of *rpa1* deletion mutants**

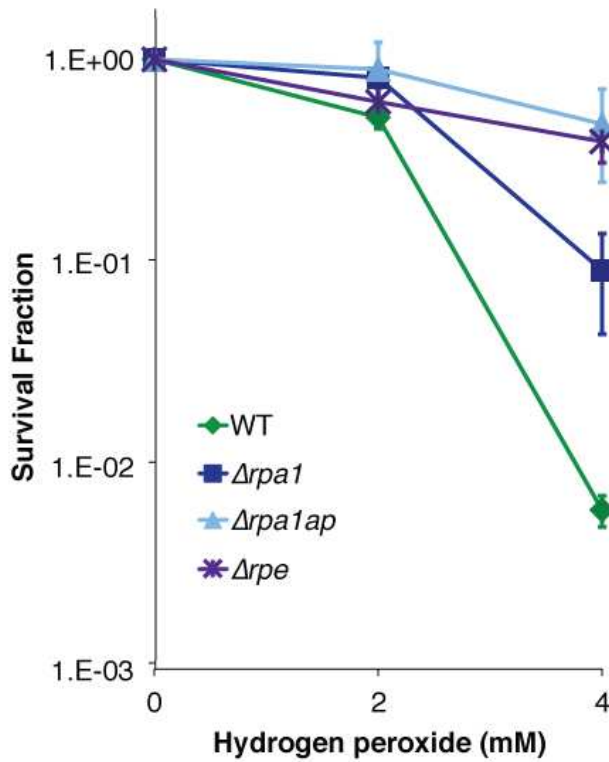
The survival fraction of wild-type (WT) H195, H1217 ( $\Delta rpa1$ ), H1216 ( $\Delta rpa1ap$ ), H1134 ( $\Delta rpe$ ) and H1246 ( $\Delta rpa1$  operon). The data shown is a mean and standard error of three repeats.

#### **4.4.3. Hydrogen peroxide DNA damage assays**

Oxidative damage occurs naturally in the cell due to the presence of reactive oxygen species (ROS), which are constantly generated as by-products of aerobic metabolism. Hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) induces oxidative DNA damage that can result in single-stranded DNA breaks (Freidberg, Walker *et al.* 2006).

To examine if RPA1, RPA1AP and RPE function in these pathways, single deletion mutants were analysed, in parallel with H195 (*rpa1*<sup>+</sup>, *rpa1ap*<sup>+</sup>, *rpe*<sup>+</sup>), in H<sub>2</sub>O<sub>2</sub> assay, shown in Figure 4.28. Liquid culture aliquots were incubated for an hour with 10 µl of diluted H<sub>2</sub>O<sub>2</sub>, giving final concentrations of 2, 4 and 8 mM, before being plated out.

The data shows all of the single deletion mutants are slightly resistant to H<sub>2</sub>O<sub>2</sub>, this could be explained by the slow growth rate shown in Figure 4.29, providing more time for the repair of oxidative damage induced by H<sub>2</sub>O<sub>2</sub>.

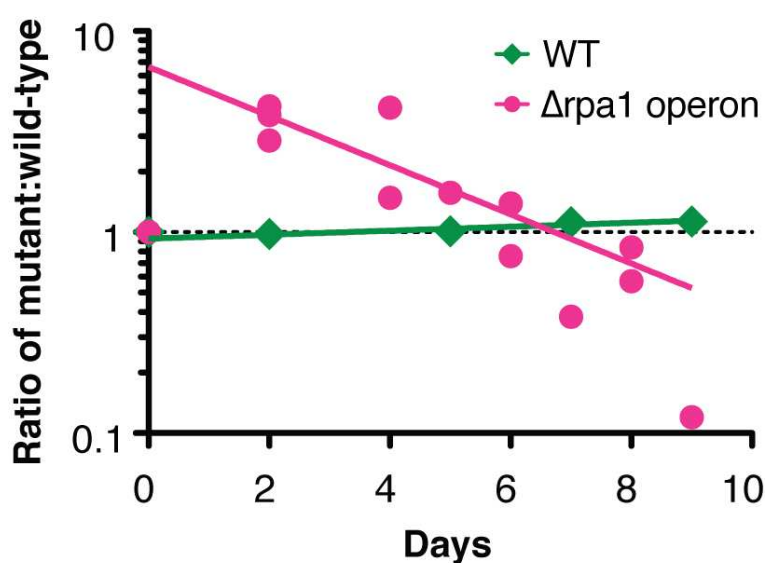


**Figure 4.28 H<sub>2</sub>O<sub>2</sub> sensitivity of *rpa1* deletion mutants**

The survival fraction of wild-type (WT) H195, H1217 ( $\Delta rpa1$ ), H1216 ( $\Delta rpa1ap$ ) and H1134 ( $\Delta rpe$ ). The data shown is a mean and standard error of three repeats.

## 4.5. Growth competition assay

To examine if the  $\Delta rpa1$  operon mutant has a growth defect it was grown competitively with a H1040 using the  $\beta$ -galactosidase ( $\beta$ -gal), encoded by *bgaH* (Holmes and Dyll-Smith 1990), to distinguish between the two strains. This type of growth analysis was chosen because competitive growth provides high sensitivity in identifying minor growth defects that would be unidentifiable in a traditional growth assay, as shown with  $\Delta mre11 \Delta rad50$  mutants in *H. volcanii* (Delmas, Shunburne *et al.* 2009). When plated on Hv-YPC (+*thy*) agar and sprayed with X-gal, H1040 is able to metabolise the X-gal due to the presence of  $\beta$ -gal and will consequently turn blue, while the  $\Delta rpa1$  operon mutant is unable to metabolise the X-gal due to the absence of  $\beta$ -gal and will remain red. To confirm there was no advantage to a strain possessing *bgaH*, a growth competition assay using H195 ( $\beta$ -gal -) and H1040 ( $\beta$ -gal +) was carried out in parallel. Figure 4.29 shows the  $\Delta rpa1$  operon deletion mutant has a small growth defect of 2.73% relative to the wild type.



**Figure 4.29** Competition assay of  $\Delta rpa1$  operon vs wild-type

The ratio of mutant:wild-type (WT) of H1246 ( $\Delta rpa1$  operon,  $\beta$ -gal-) to H1040 ( $\beta$ -gal+) shown in pink, and the control H195 ( $\beta$ -gal -) to H1040 ( $\beta$ -gal +) is shown over a time course. The curve fittings of the trend lines, drawn between the three repeats of data sets, was based on the time period of 80 generations and a linear regression curve fit.

## 4.6. Discussion

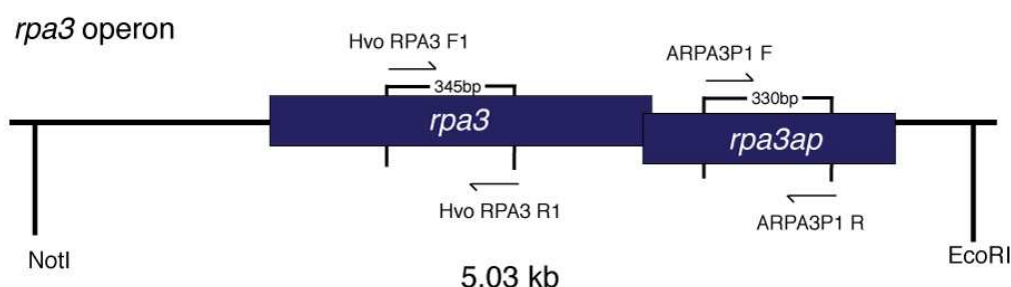
The ease with which genomic deletions of *rpal*, *rpalap*, *rpe* and *rpal* operon deletion were made indicates that Rpa1 and the associated proteins are not essential for cell survival. However, they may still have a role in DNA repair and recombination.

The DNA damage assays indicate *rpal*, *rpalap* and *rpe* do not play a significant role in DNA repair. The slight growth defect of the *rpal* operon deletion mutant demonstrates that the products of the *rpal* operon aid cell growth and duplication but are not essential. However, an element of redundancy between RPA1 and RPA3 or RPA2 could explain the lack of a significant growth defect and DNA damage phenotype. Attempting to generate a double  $\Delta$ rpal operon  $\Delta$ rpap3 operon and a double  $\Delta$ rpap1 operon  $\Delta$ rpap2 deletion mutant could test this hypothesis.

## Chapter 5: RPA3

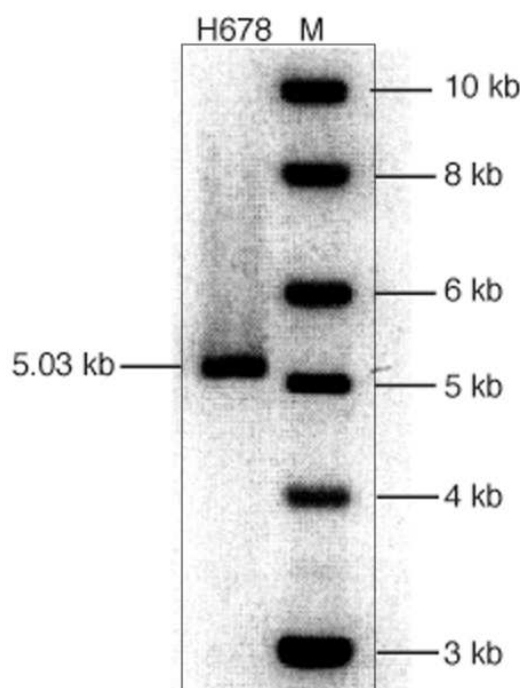
### 5.1. Isolation of the *rpa3* operon

The genomic sequence of *H. volcanii* was used to identify appropriate restriction sites to obtain a genomic clone of *rpa3*. Genomic DNA from *H. volcanii* strain was digested with *NotI* and *EcoRI* to give a predicted fragment size of ~5 kb, shown in Figure 5.1, followed by Southern hybridisation shown in Figure 5.2.



**Figure 5.1** *NotI* and *EcoRI* genomic digest to isolate *rpa3* operon

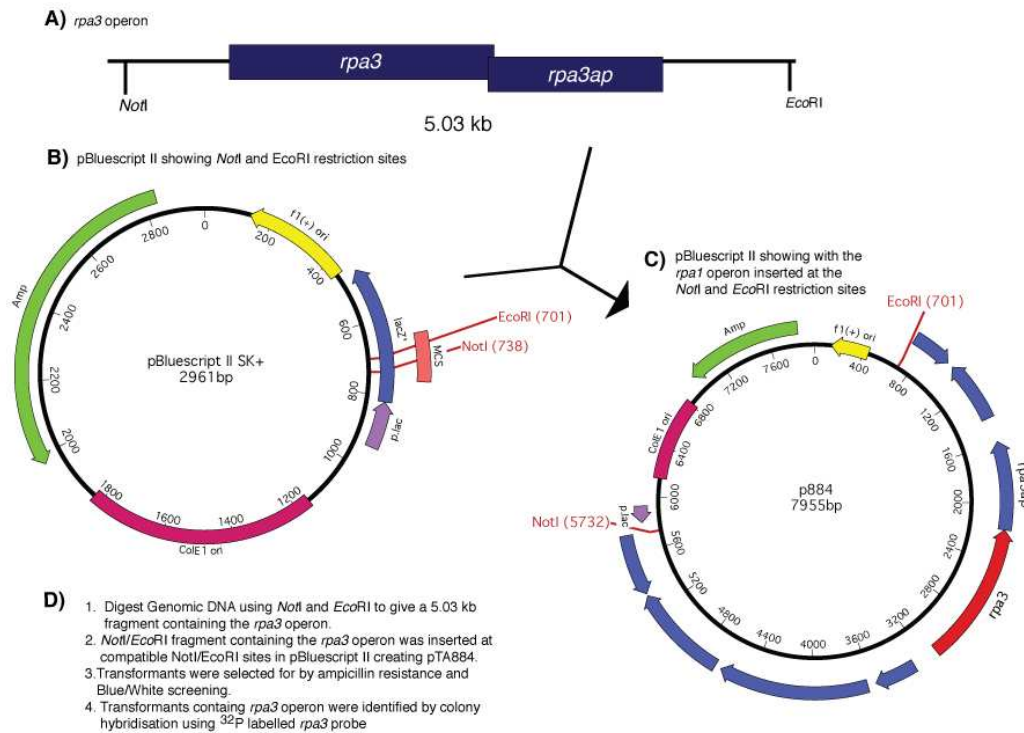
The location of *NotI* and *EcoRI* restriction sites generating a ~5 kb fragment are shown, as well as the primers used to generate the 345 bp and 330bp products that were used in the *rpa3* and *rpa3ap* colony hybridisations respectively. The associated gene containing an OB-fold (*rpa3ap*) is shown.



**Figure 5.2** Southern blot analysis of *rpa3* genomic fragment

Lane M shows the size marker lane. Genomic DNA (H678) was digested with *NotI* and *EcoRI*, and probed with *rpa3* PCR product, shown in Figure 5.1.

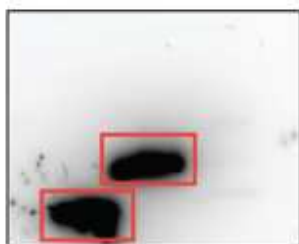
A library of *NotI*/*EcoRI* DNA fragments 4.5-5.5 kb in size was cloned in the plasmid pBluescript II SK+, which features an ampicillin marker, as shown in Figure 5.3.



**Figure 5.3 Cloning strategy to isolate the *rpa3* operon**

A) *NotI* and *EcoRI* restriction sites used to insert the *rpa3* operon into B) pBluescript II to generate C) a genomic clone of the *rpa3* operon, pTA884.

The plasmid library was transformed into *E. coli* XL-1, blue/white screening was used to isolate transformants with an insert. These transformants were then screened by colony hybridisation for the presence of *rpa3* using an *rpa3* probe labelled with <sup>32</sup>P, as shown in Figure 5.4. Plasmid extraction followed by sequencing was used to confirm the isolation of the *rpa3* operon in pTA884.



**Figure 5.4. Colony hybridisation screening for *rpa3* operon.**

Screening of transformants containing *rpa3* operon by colony hybridisation with  $^{32}\text{P}$  labelled *rpa3* probe (See figure 5.3).

## 5.2 Generation of single *rpa3* and *rpa3ap* deletion mutants

Upstream and downstream regions of the gene to be deleted were amplified, introducing novel restriction sites to allow ligation of the two regions into the plasmid pTA131, as described previously. The deletion construct was then used to generate a genomic deletion in *H. volcanii* using the counter-selective pop-in/pop-out method, shown in Figure 4.6 (Allers, Ngo *et al.* 2004).

### 5.2.1 Construction of *rpa3* deletion mutant

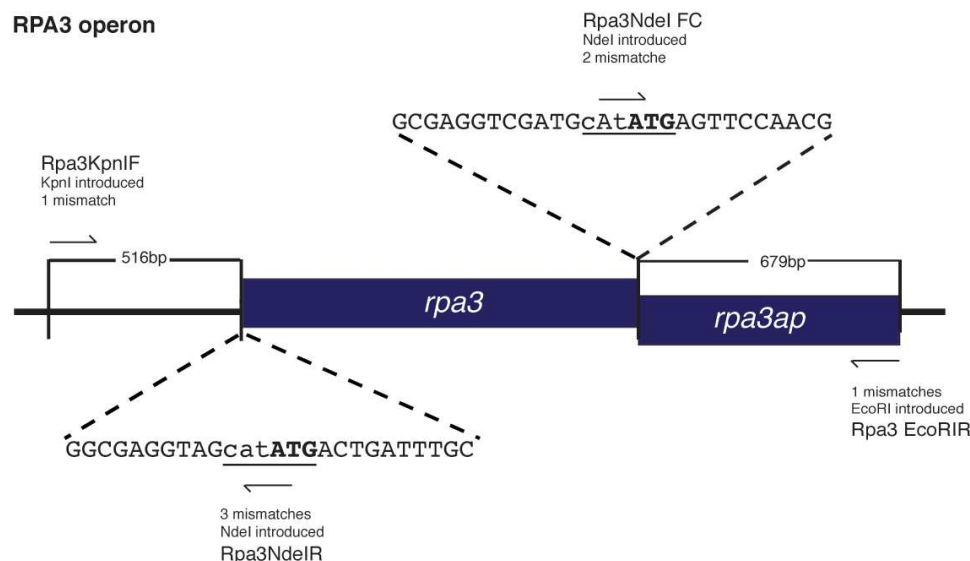
#### Design of primers

Upstream and downstream regions of *rpa3* were amplified from the genomic *rpa3* clone pTA884 by touchdown PCR using the primers shown in Table 5.1, introducing external *KpnI* and *EcoRI* restriction sites and an internal *NdeI* restriction site.

Primer	Sequence (5'-3')	Site inserted
Rpa3NdeI FC(DS)	GCGAGGTCGATGcAt <b>ATG</b> AGTTCCAACG	<i>NdeI</i>
Rpa3EcoRIR(DS)	GACGGTGGAAtTCGGCCGTCG	<i>EcoRI</i>
Rpa3NdeIR(US)	GCAAATCAGTcat <b>ATG</b> CTACCTCGCC	<i>NdeI</i>
Rpa3KpnI F(US)	GCCGGTGGTaCCACAGCCTC	<i>KpnI</i>

**Table 5.1 Design of deletion primers for *rpa3***

The sequence for each deletion primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of the upstream (US) and downstream (DS) regions. Start codon of the *rpa3* and *rpa3* are shown in bold in primers Rpa3NdeIR and Rpa3NdeI FC respectively.



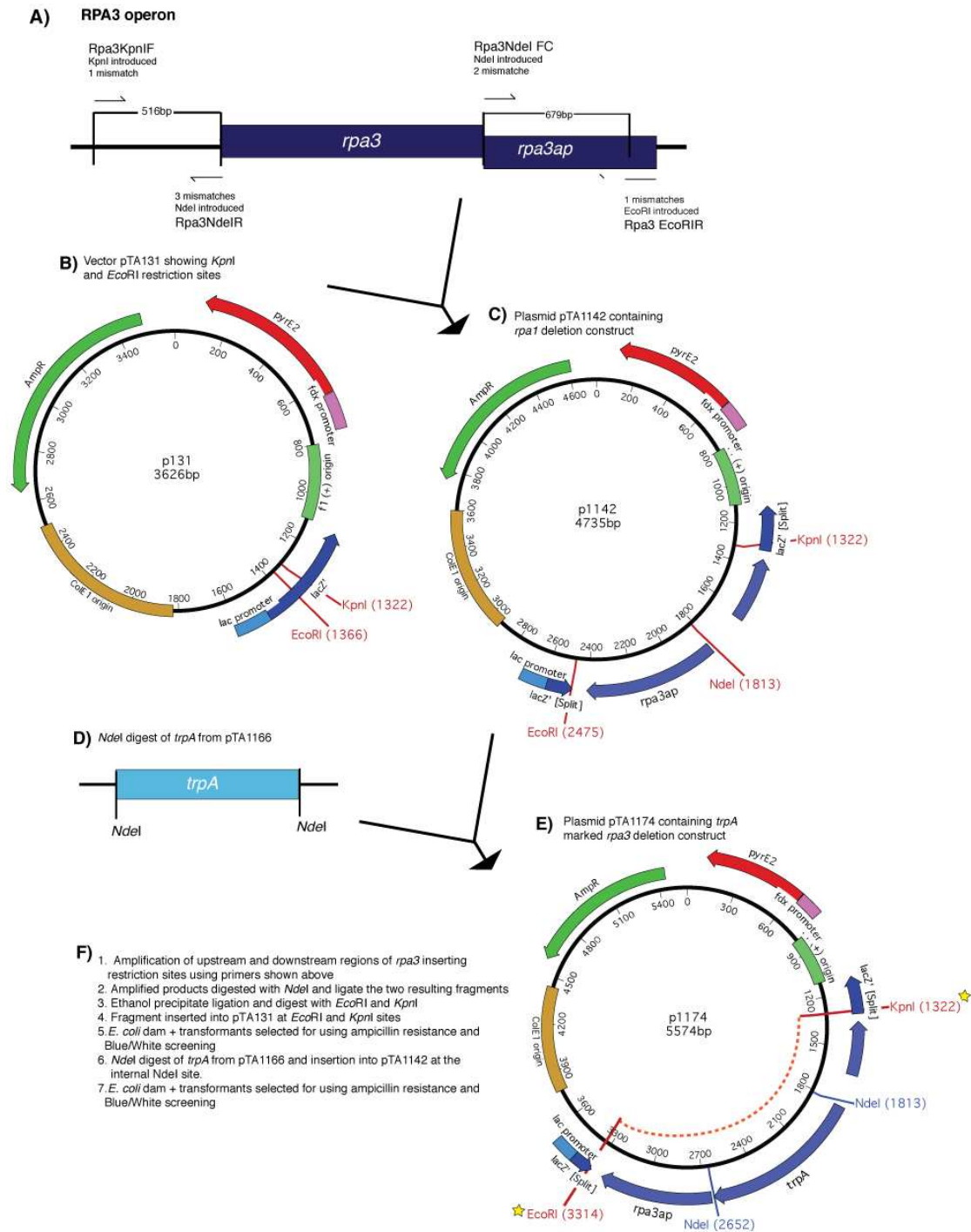
**Figure 5.5 Introduction of internal *NdeI* sites to delete *rpa3***

The introduction of internal *NdeI* restriction sites maintains the reading frame and thus the start codon of the downstream *rpa3ap*.

### Construction of the deletion plasmid

The PCR products of the upstream and downstream regions were gel purified, digested with *NdeI* and ligated. The upstream/downstream ligation was ligated into the plasmid pTA131 using the *KpnI* and *EcoRI* sites, and was transformed into *E. coli* XL-1 strain, as shown in Figure 5.6. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct construct, pTA1142. The *trpA* marker was then digested from pTA1166 (see Figure 4.8 for details) using *NdeI* and inserted at the internal *NdeI* site of pTA1142. Restriction digest of plasmid extractions from transformants was used to screen for the correct construct, pTA1174.



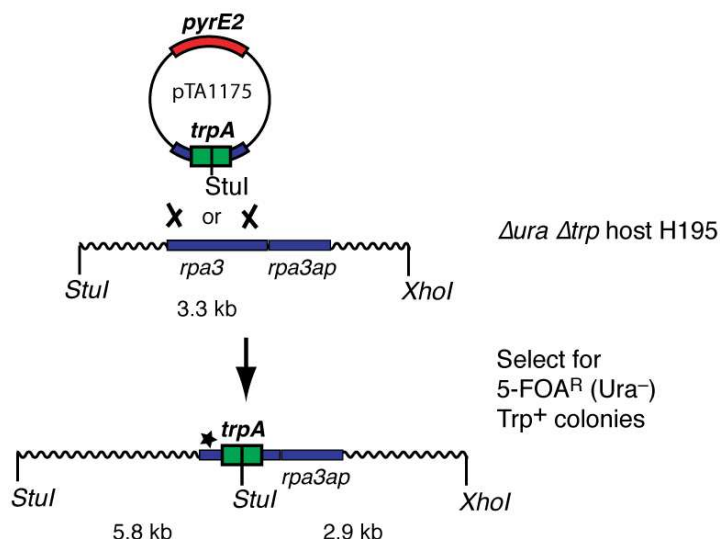


**Figure 5.6 Construction of the *rpa3* deletion**

Cloning technique used to generate the deletion construct of *rpa3*, the binding sites of the primers are shown A) the restriction site in the vector used to insert the upstream and downstream ligation B) and the resulting deletion construct pTA1142 C). The *trpA* marker was digested with *NdeI* from pTA1166 D) and inserted into pTA1142 at the internal *NdeI* site generating pTA1174 E). A basic sequence of construction is shown in F). Orange dotted line and yellow stars denotes region that was used as a  $^{32}\text{P}$  labelled probe in Southern hybridisation shown in Figure 5.9.

### Pop-in/ pop-out of the deletion construct

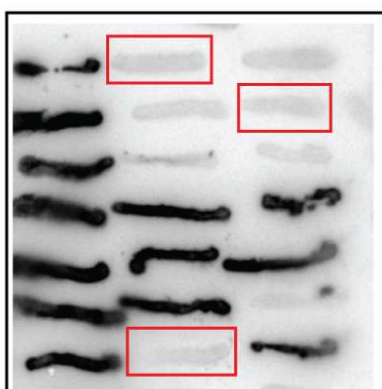
Transformation into the *E. coli dam* mutant generated the plasmid pTA1175, which was transformed into *H. volcanii* generating the pop-in H1200 to generate a genomic deletion of *rpa3*, as shown in Figure 5.7.



**Figure 5.7. Schematic diagram of *rpa3* deletion construction**

The deletion plasmid pTA1175 was constructed as described in Figure 5.6. Restriction digest with *XhoI* and *StuI* and the resulting fragments expected that were used in the Southern blot analysis are shown. Integration onto the genome by homologous recombination (pop-in), generating H1200, followed by loss of the plasmid by homologous recombination (pop-out) whilst selecting for *trpA*<sup>+</sup> resulted in the  $\Delta rpa3$  strain (H1244). The region that the <sup>32</sup>P probe hybridises to is denoted by a star.

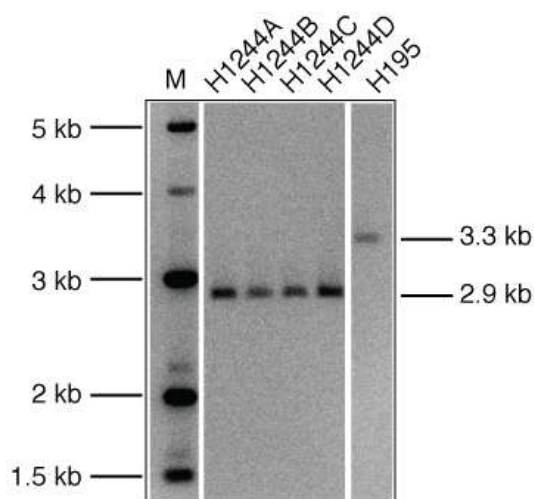
Potential deletions were screened by colony hybridisation, using a <sup>32</sup>P labelled *rpa3* probe (see Figure 5.1), pop-outs that did not hybridise to the probe were selected as potential deletion mutants, as shown in Figure 5.8.



**Figure 5.8 Colony lift and hybridisation of H1200 popouts**

Popouts (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv-Ca +thy +ura, to maintain selection for *trpA* marked deletion construct in case of merodiploidy. Three popout candidates that failed to hybridise to the *rpa3* probe are indicated by red boxes. The genomic sequence of *H. volcanii* was used to

select appropriate restriction enzymes to digest the genomic DNA of the potential deletion mutants (H1244) that was then used in Southern blotting to confirm genomic deletion of *rpa3*. A  $^{32}\text{P}$  labelled probe of the upstream, autotrophic marker and downstream region of pTA1174 was used in the hybridisation, as shown in Figure 5.6.



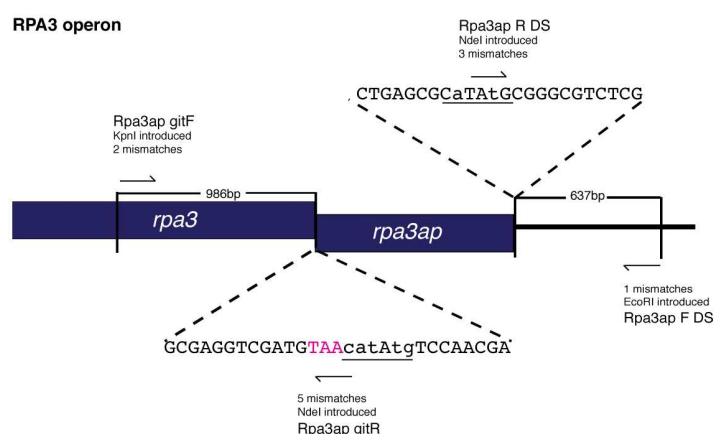
**Figure 5.9 Southern blot analysis of H1200 popouts (H1244)**

Lane M contains the size marker. Genomic DNA was digested with *Stu*I and *Xho*I and probed with the flanking regions and the *trpA* marker of the deletion construct p1174 *Eco*RI/*Kpn*I. Genomic DNA of H195 was used as a wild-type control.

## 5.2.2 Construction of *rpa3ap* deletion mutant

### Design of primers

Upstream and downstream regions of *rpa3* were amplified from the genomic *rpa3* clone pTA884 by touchdown PCR using the primers shown in Table 5.2, introducing external *Kpn*I and *Eco*RI restriction sites and an internal *Nde*I restriction site.



**Figure 5.10 Introduction of internal *Nde*I sites to delete *rpa3ap***

Introduction of internal *Nde*I restriction sites to allow insertion of the *trpA* marker and maintain the *rpa3* sequence.

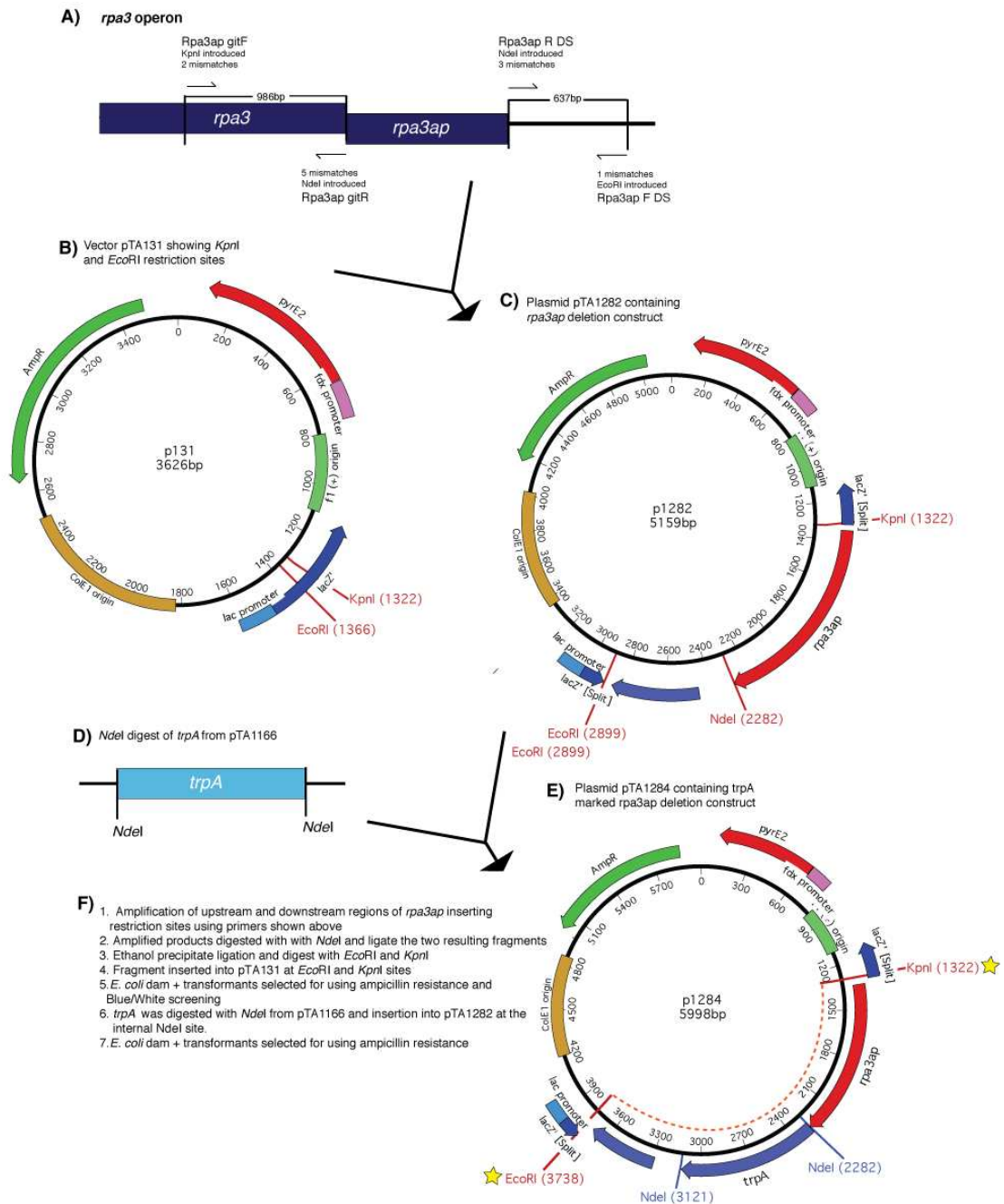
Primer	Sequence	Site inserted
Rpa3ap gitF (US)	CTCCCAATGG <u>gtACCAAGGTGGAGGC</u>	<i>KpnI</i>
Rpa3ap gitR (US)	GCGAGGTCGATG <u>TAA</u> catAtgTCCAACGA	<i>NdeI</i>
Rpa3ap R DS	CTGAGCGC <u>CaTA</u> tGCGGGCGTCTCG	<i>NdeI</i>
Rpa3ap F DS	CTCGCTGAA <u>tTCGGTGGGTGC</u>	<i>EcoRI</i>

**Table 5.2 Primers for the construction the *rpa3ap* deletion construct**

Primers used to generate the RPA3ap deletion construct are shown. The sequence for each deletion primer is shown with the base changes (in lower case) introducing a novel restriction site (underlined) to allow cloning of the upstream (us) and downstream (ds) regions. The stop codon for *rpa3* is shown in blue.

### Construction of the deletion plasmid

The PCR products of the upstream and downstream regions were gel purified, digested with *NdeI* and ligated together. The upstream/downstream ligation was ligated into the plasmid pTA131 using the *KpnI* and *EcoRI* sites, and was transformed into *E. coli* XL-1 strain, as shown in Figure 5.11. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct construct, pTA1282. The *trpA* marker was then digested from pTA1166 using *NdeI* and inserted at the internal *NdeI* site in pTA1282, followed by transformation into *E. coli* XL-1, as shown in Figure 5.11. Restriction digest of plasmid extractions from transformants was used to screen for the correct construct, pTA1284.

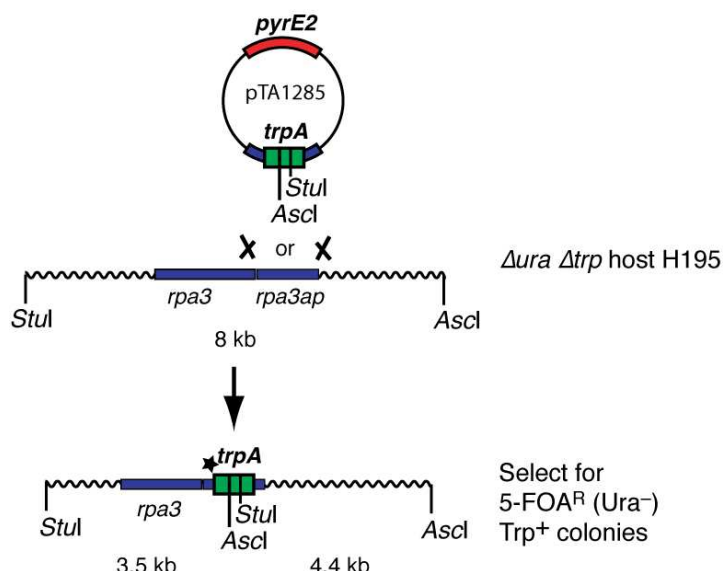


**Figure 5.11 construction of the *rpa3ap* deletion**

Cloning technique used to generate the deletion construct of *rpa3ap*, the binding sites of the primers are shown A) the restriction site in the vector used to insert the upstream and downstream ligation B) and the resulting deletion construct pTA1282 C). The *trpA* marker was digested with *NdeI* from pTA1166 D) and inserted into pTA1282 at the internal *NdeI* site generating pTA1284 E). A basic sequence of construction is shown in F). Orange dotted line and yellow stars denotes region that was used as a  $^{32}\text{P}$  labelled probe in Southern hybridisation shown in Figure 5.14.

### Pop-in/ pop-out of the deletion construct

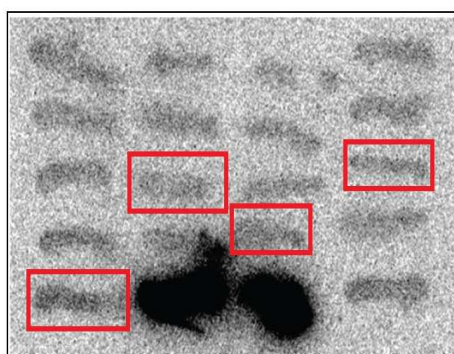
Following transformation into the *E. coli dam*- the plasmid pTA1185 was transformed into *H. volcanii* H195 generating H1409, to generate a genomic deletion of *rpa3ap* in strain H1410, as shown in Figure 5.12.



**Figure 5.12 Schematic diagram of *rpa3ap* deletion construction**

The deletion plasmid pTA1171 was constructed as described in Figure 5.11. Restriction digest with *AscI* and *StuI* (see text) and the resulting fragments expected are shown. Integration into the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) whilst selecting for *trpA*<sup>+</sup> resulted in the  $\Delta rpa3ap$  strain. The region that the <sup>32</sup>P labelled probe hybridises to is denoted by a star.

Potential deletions were screened by colony hybridisation, using a <sup>32</sup>P labelled *rpa3ap* probe (see Figure 5.1), pop-outs that did not hybridise to the probe were selected as potential deletion mutants, as shown in Figure 5.13.

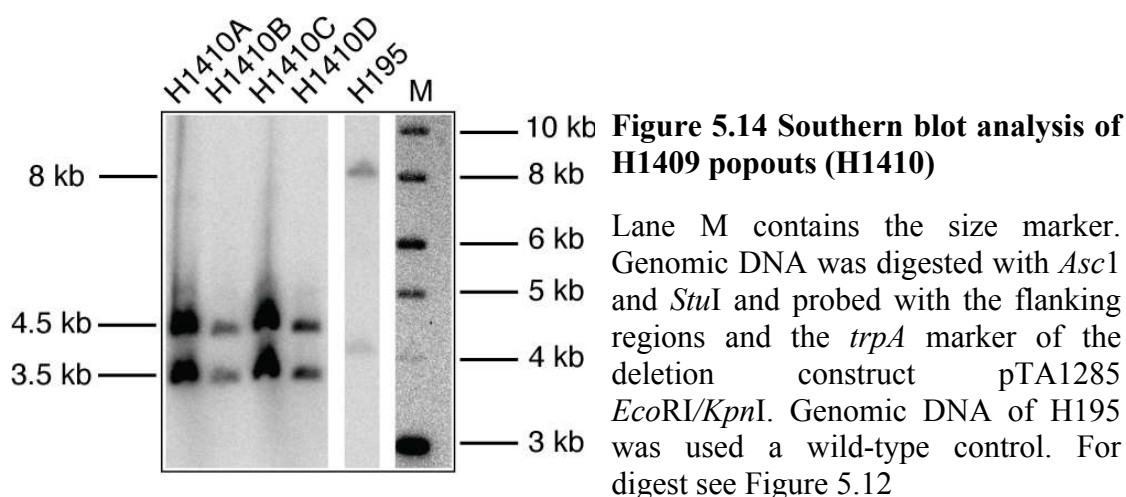


**Figure 5.13 Colony lift and hybridisation of H1409 popouts**

Popouts (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv-Ca +thy +ura, maintaining selection for *trpA* marked deletion construct. Three popout candidates that failed to hybridise to the *rpa3ap* probe are indicated by red boxes.



The genome sequence of *H. volcanii* was used to determine an appropriate restriction digest of the potential deletion mutants, *AscI* and *StuI*, which was then used in Southern blotting to confirm genomic deletion of *rpa3ap*. A  $^{32}\text{P}$  labelled probe of the upstream, *trpA* marker and downstream region of pTA1284 was used in the hybridisation, as shown in Figure 5.11.

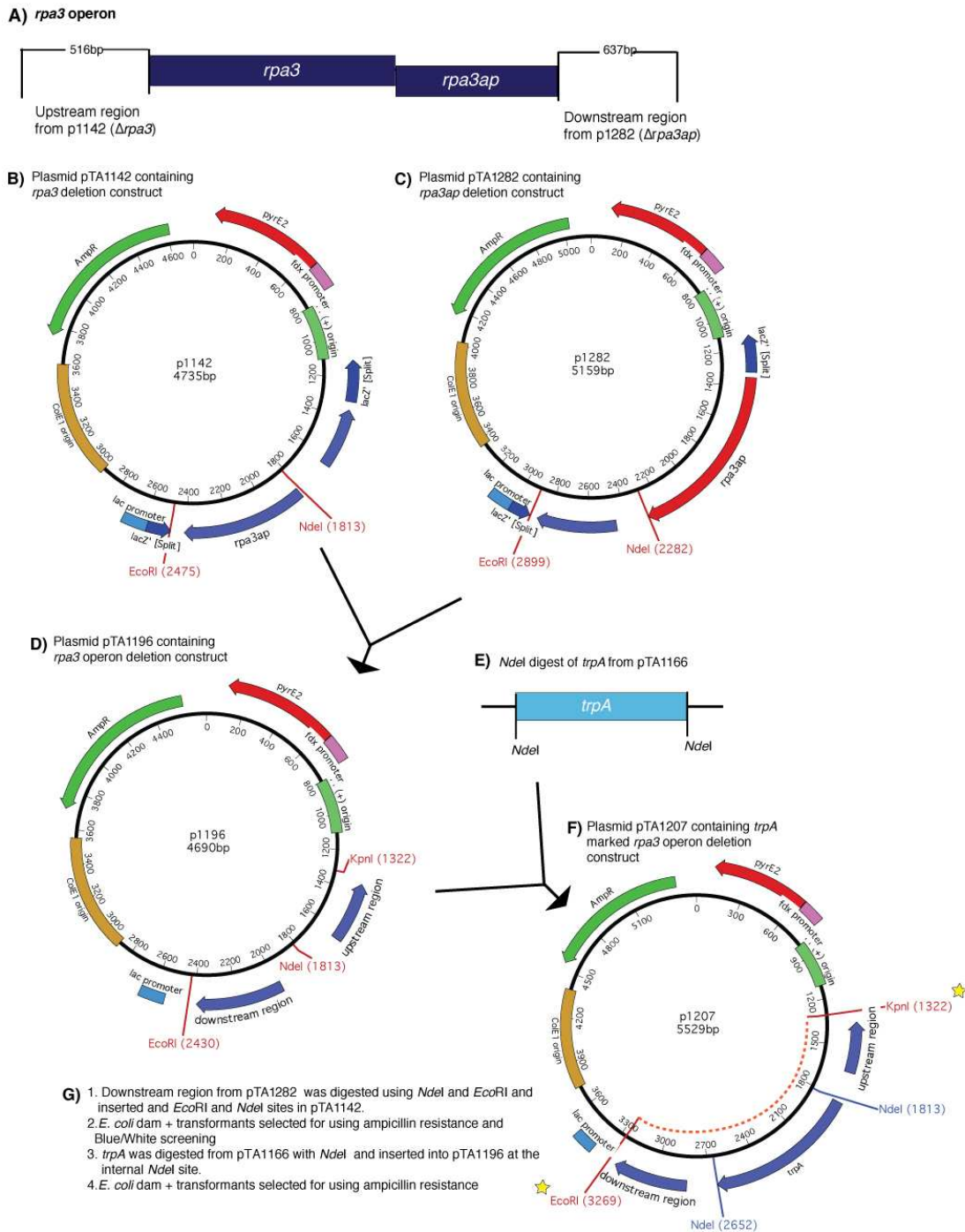


## 5.3 Generation of *rpa3* operon deletion

### 5.3.1 Construction of *rpa3* operon deletion mutant

#### Construction of the deletion plasmid

The downstream fragment of the *rpa3* operon was digested from the *rpa3ap* deletion construct pTA1282 using *EcoRI* and *NdeI*. This was inserted into the *rpa3* deletion construct pTA1142, replacing the existing downstream fragment, as shown in Figure 5.15. The *rpa3* operon deletion construct was transformed into *E. coli* XL-1, and restriction digest of plasmid extractions from transformants was used to screen for the correct construct, pTA1196. The *trpA* marker was then digested from pTA1166 using *NdeI* and inserted at the internal *NdeI* site in pTA1196, followed by transformation into *E. coli* XL-1, as shown in Figure 5.15. Restriction digest of plasmid extractions from transformants was used to screen for the correct construct, pTA1207.



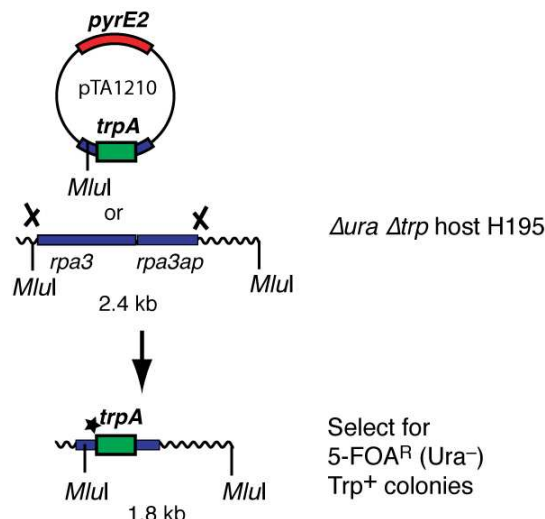
**Figure 5.15. Construction of the *rpa3* operon deletion**

Figure shows the cloning technique used to generate the deletion construct of *rpa3* operon, including the upstream and downstream fragments used A) the restriction sites in the pTA1142 used to replace the downstream fragment B) and the restriction sites used to digest the downstream fragment from pTA1282 C). Generating pTA1196 D), *trpA* was digested from pTA1166 E) and inserted into pTA1196 at an internal *NdeI* site, generating pTA1207 F) A basic sequence of construction is shown in G) Orange dotted line and yellow stars denote a region that was used as a <sup>32</sup>P labelled probe in Southern hybridisation shown in Figure 5.18.



### Pop-in/ pop-out of the deletion construct

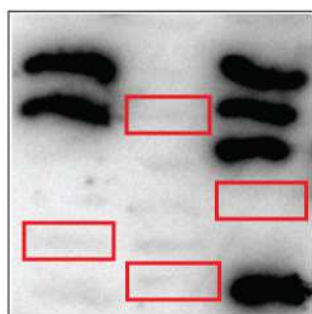
Following transformation into the *E. coli dam*- the plasmid pTA1210 was transformed into *H. volcanii* H195 generating the pop-in H1257, which was used to generate a genomic deletion of *rpa3* operon (H1260), as shown in Figure 5.16.



**Figure 5.16 Schematic diagram of *rpa3* operon deletion construction**

The deletion plasmid pTA1210 was constructed as described in Figure 5.15. Restriction digest with *MluI* and the resulting fragments expected are shown. Integration onto the genome by homologous recombination (pop-in), H1257, followed by loss of the plasmid by homologous recombination (pop-out) resulted in the  $\Delta rpa3$  operon strain (H1260). The region that the <sup>32</sup>P labelled probe hybridises to is denoted by a star.

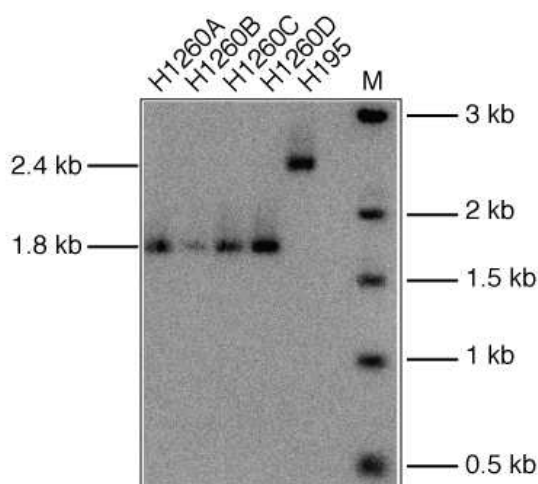
Potential deletions were screened by colony hybridisation using <sup>32</sup>P labelled *rpa3* probe, pop-outs that did not hybridise to the probe were selected as potential deletion mutants, as shown in Figure 5.17.



**Figure 5.17 Colony lift and hybridisation of H1257 popouts**

Pop-outs (*ura*-) selected for on 5FOA were patched out on Hv- YPC. Four pop-out candidates that failed to hybridise to the *rpa3* probe are indicated by red boxes.

The genomic sequence of *H. volcanii* was used to determine an appropriate restriction digest, *Mlu*I to screen the potential deletion mutants, this was then used in a Southern blotting to confirm genomic deletion of *rpa3* operon. A <sup>32</sup>P labelled probe of the upstream, *trpA* and downstream region of pTA1207 was used in the hybridisation, as shown in Figure 5.15.



**Figure 5.18. Southern blot analysis of H1257 popouts (H1260)**

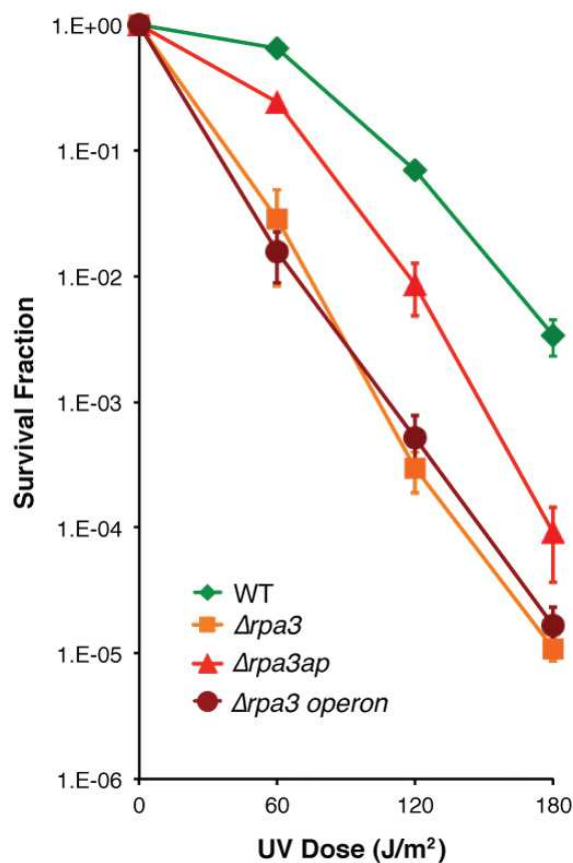
Lane M contains the size marker. Genomic DNA was digested with *Mlu*I and probed with the flanking regions of the deletion construct pTA1207 digested with *Eco*RI/*Xba*I. Genomic DNA of H195 was used as a wild-type control. See Figure 5.16 for restriction digest details.

## 5.4 DNA damage assays of deletion mutants

To examine if RPA3 or RPA3AP play a role in DNA repair, DNA damage assays were performed in the same manner as had been carried out on the *rpa1* deletion mutants.

### 5.4.1 UV DNA damage assays

To examine if RPA3 and RPA3AP function in the repair of UV-induced DNA damage, single deletion mutants and the *rpa3* operon deletion mutant were analysed in parallel with a H195 (*rpa3*<sup>+</sup>, *rpa3ap*<sup>+</sup>) in a UV assay, shown in Figure 5.19. The data shows that both of the single deletion mutants and the *rpa3* operon mutants are significantly sensitive to UV, the *rpa3ap* deletion is the least sensitive of the three. This indicates that RPA3 and RPA3AP are involved significantly in the removal and repair of UV DNA damage, most probably by the same pathway.

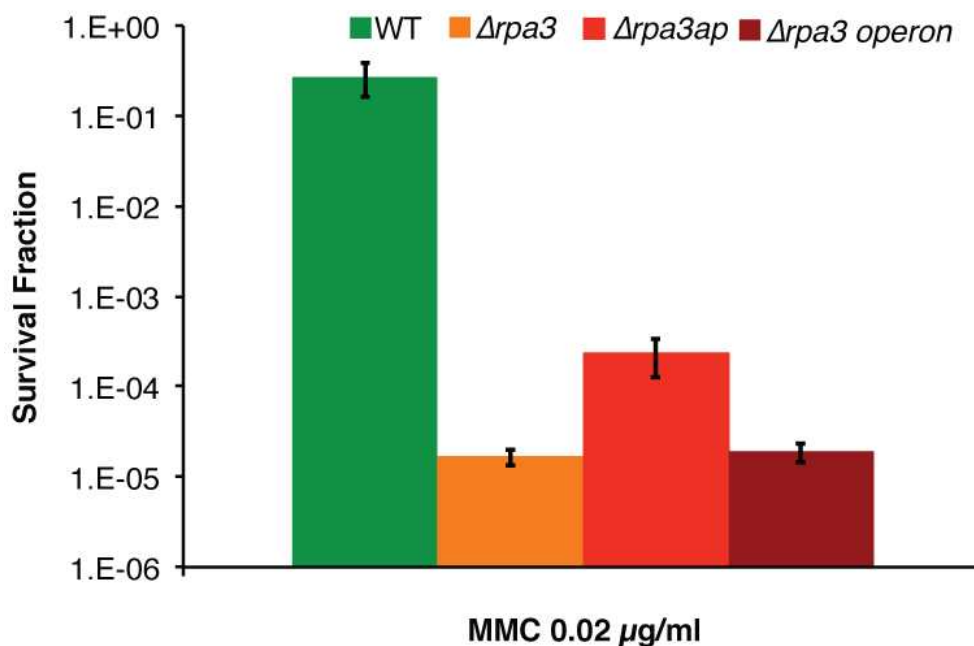


**Figure 5.19 UV sensitivity of *rpa3* deletion mutants**

The survival fraction of wild-type (WT) H195, H1244 ( $\Delta rpa3$ ), H1410 ( $\Delta rpa3ap$ ) and H1260 ( $\Delta rpa3$  operon). The data shown is a mean average of three and standard error of three repeats.

#### 5.4.2 MMC DNA damage assays

To examine if RPA3 and RPA3AP function in the repair of MMC-induced DNA damage, single deletion mutants and the *rpa3* operon deletion mutant were analysed in MMC assay, shown in Figure 5.20. Hv-YPC +thy plates containing 0.02  $\mu\text{g/ml}$  MMC were used in parallel with a H195 (*rpa3*+, *rpa3ap*+). The data shows the single deletion mutants and the *rpa3* operon mutant are significantly sensitive to MMC, the *rpa3ap* deletion mutant is the least sensitive of the deletion mutants. This indicates that both RPA3 and RPA3AP are involved significantly in the removal and repair of ICLs, most probably by the same pathway.

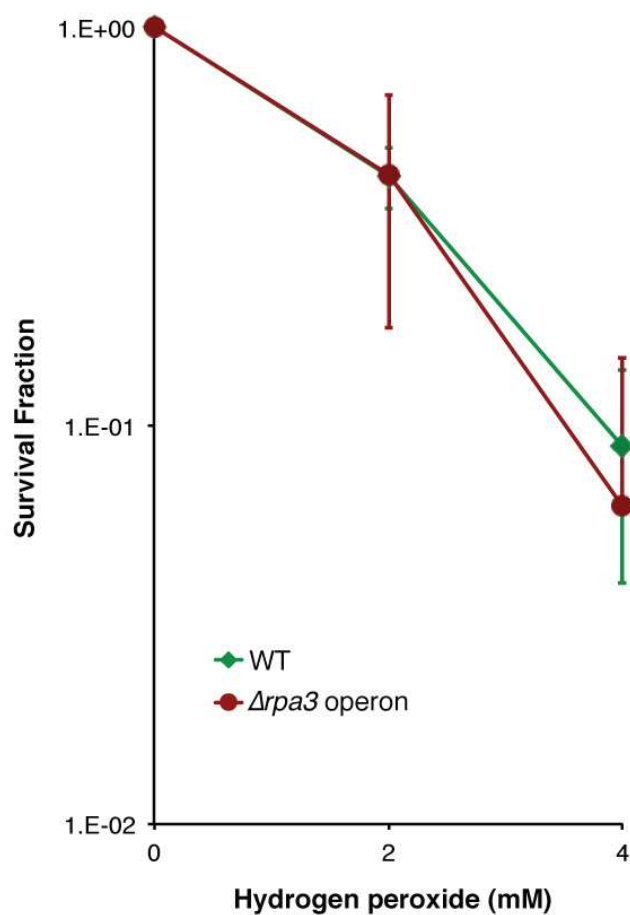


**Figure 5.20 MMC sensitivity of *rpa3* deletion mutants**

The survival fraction of wild-type (WT) H195, H1244 ( $\Delta rpa3$ ), H1410 ( $\Delta rpa3ap$ ), and H1260 ( $\Delta rpa3$  operon). The data shown is a mean and standard error of three repeats.

#### 5.4.3 Hydrogen peroxide DNA damage assays

To examine if RPA3 and RPA3AP function in the repair of oxidative damage, induced by  $H_2O_2$ , the *rpa3* operon deletion mutant was analysed, in parallel with a H195 (*rpa3*<sup>+</sup>, *rpa3ap*<sup>+</sup>), in a  $H_2O_2$  assay as shown in Figure 5.21 The data shows that the *rpa3* operon deletion mutant is not sensitive to  $H_2O_2$ , suggesting that RPA3 and RPA3AP are not involved significantly in the removal and repair of oxidative DNA damage.

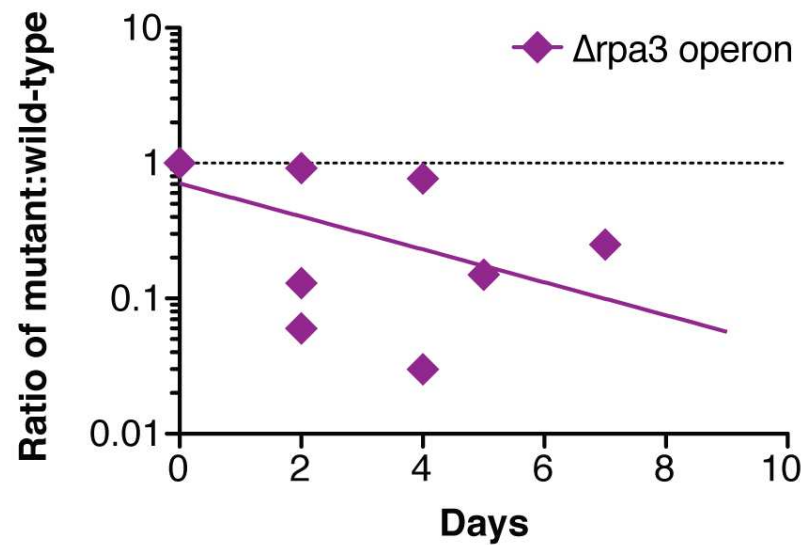


**Figure 5.21 H<sub>2</sub>O<sub>2</sub> sensitivity of *rpa3* deletion mutants**

The survival fraction of wild-type (WT) H195 and ( $\Delta rpa3$  operon) H1260. The data shown is a mean average of three repeats and error bars indicate standard error.

## 5.5 Growth competition assay

To examine if the  $\Delta rpa3$  operon mutant has a growth defect, it was grown competitively with H1040 ( $rpa3^+$ ,  $rpa3ap^+$ ,  $rpe^+$ ,  $\beta\text{-gal}^+$ ) in a growth competition assay, as described in Chapter 4. The  $\Delta rpa3$  operon deletion mutant showed a small growth defect of 4% relative to the wild type.



**Figure 5.22 Competition assay of  $\Delta rpa3$  operon vs. wild-type**

The ratio of mutant:wild-type (WT) of H1260 ( $\Delta rpa3$  operon,  $\beta\text{-gal}^-$ ) to H1040 ( $\beta\text{-gal}^+$ ) is shown in purple. The curve fittings of the trend lines, drawn between the three repeats of data sets, were fitted based on the time period of 80 generations and a linear regression curve fit.

## 5.6 Discussion

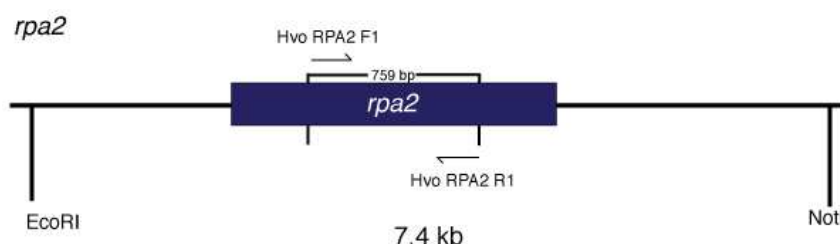
The ability to make genomic deletions of *rpa3*, *rpa3ap* and the *rpa3* operon indicates that RPA3 and its associated protein are not essential for cell survival. However, it is clear they have a significant role in a variety of DNA repair pathways, as shown by the DNA damage assays. The sensitivity of the *rpa3*, *rpa3ap* and *rpa3* operon deletion mutants to DNA damage, and the lack of such sensitivity of the *rpa1*, *rpa1ap* and *rpa1* operon deletion mutants, suggests that in spite of the close phylogenetic relationship between RPA1 and RPA3, these two proteins have different cellular roles. The slight growth defect of the *rpa3* operon deletion mutant demonstrates that the *rpa3* operon aids cell growth and duplication but is not essential.

The DNA damage phenotype demonstrates the significance of RPA3 and RPA3AP in DNA repair, but the cells are still viable and consequently an element of redundancy between RPA3 and RPA1 or RPA2 must exist. This hypothesis will be tested by attempting to generate a double  $\Delta rpa1$  operon  $\Delta rpa3$  operon mutant, and a double  $\Delta rpa3$  operon  $\Delta rpa2$  deletion mutant.

## Chapter 6: RPA2

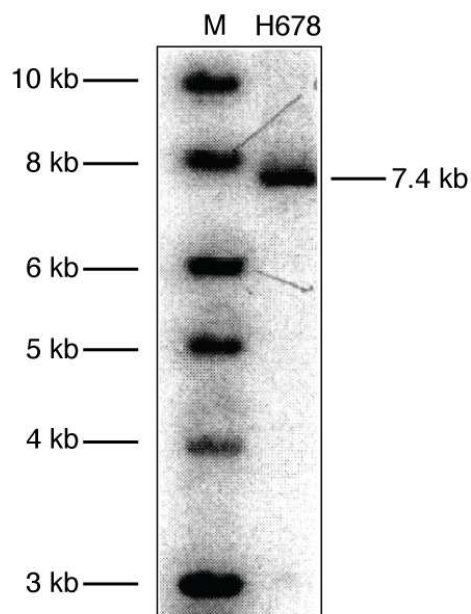
### 6.1 Isolation of the *rpa2* operon

The genomic sequence of *H. volcanii* was used to identify appropriate restriction sites to obtain a genomic clone of *rpa2*. Genomic DNA from *H. volcanii* wild-type strain H678 was digested with *EcoRI* and *NotI* to give a predicted fragment size of 7.4 kb, shown in Figure 6.1, and was confirmed by Southern hybridisation shown in Figure 6.2



**Figure 6.1** *EcoRI* and *NotI* genomic digest to isolate *rpa2*

The location of *EcoRI* and *NcoI* restriction sites generating a 7.4 kb fragment are shown, as well as the primers used to generate the 301 bp product that was used in the Southern hybridisation.

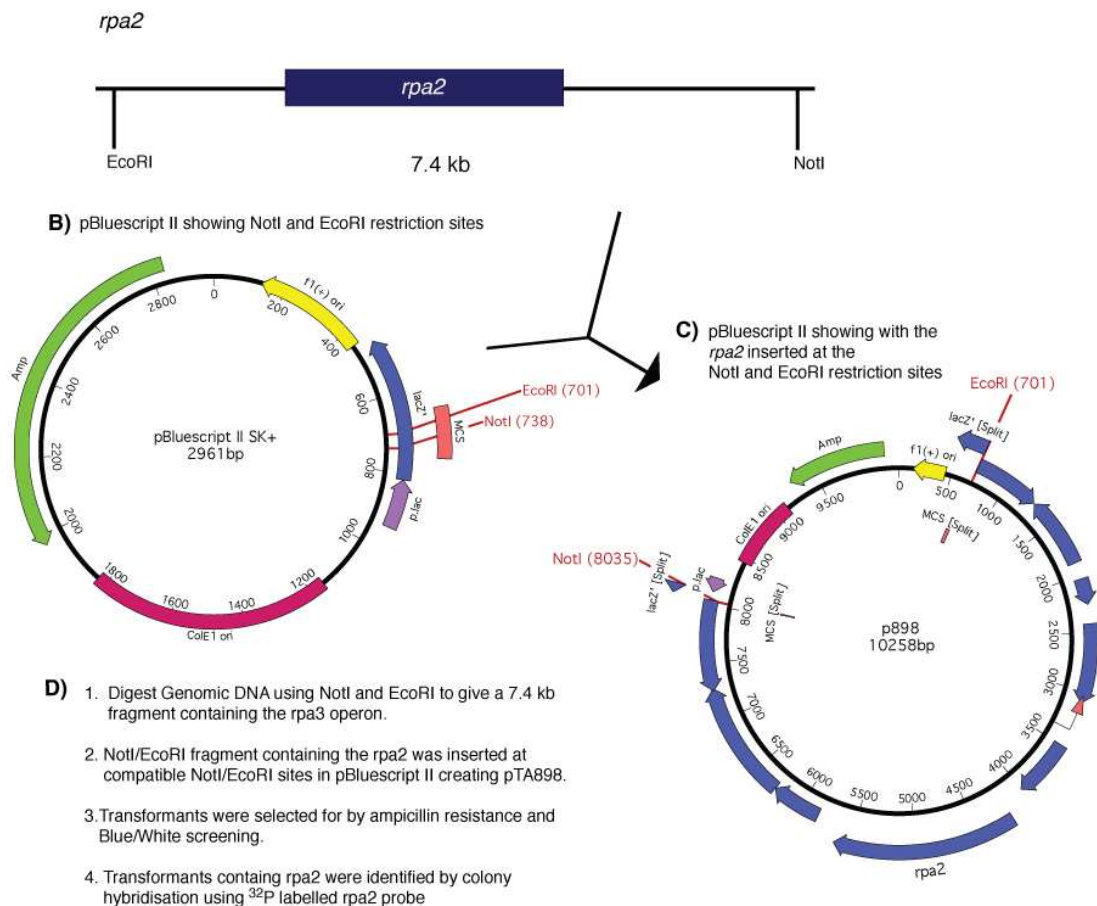


**Figure 6.2** Southern blot analysis of *rpa2* genomic clone

Lane M shows the size marker. Genomic DNA (H678) was digested with *EcoRI* and *NotI* and probed with *rpa2* PCR product, shown in Figure 6.1.1.

A library of *EcoRI/NotI* DNA fragments 7-8 kb in size was cloned in the plasmid pBluescript II SK+, which features an ampicillin marker, as shown in Figure 6.3.

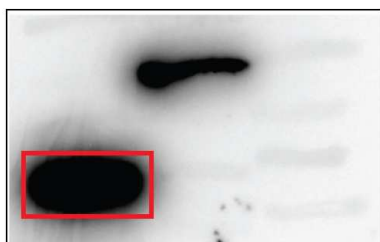




**Figure 6.3 Cloning strategy to isolate *rpa2***

A) *EcoRI/NotI* restriction sites used to insert the *rpa2* operon into B) pBluescript II at compatible *EcoRI/NotI* restriction sites, to generate C) a genomic clone of the *rpa2* operon, pTA898.

The plasmid library of *EcoRI/NotI* fragments was transformed into *E. coli* XL-1, blue/white screening was used to isolate transformants with an insert. These transformants were then screened by colony hybridisation for the presence of *rpa2* using *rpa2* probe labelled with  $^{32}\text{P}$ , as shown in Figure 6.3. Plasmid extraction followed by sequencing was used to confirm the isolation of *rpa2* in pTA898.



**Figure 6.4 Colony hybridisation screening for *rpa2* clone.**

Screening of transformants containing *rpa2* by colony hybridisation with  $^{32}\text{P}$ -labelled *rpa2* probe (See Figure 6.1.1).

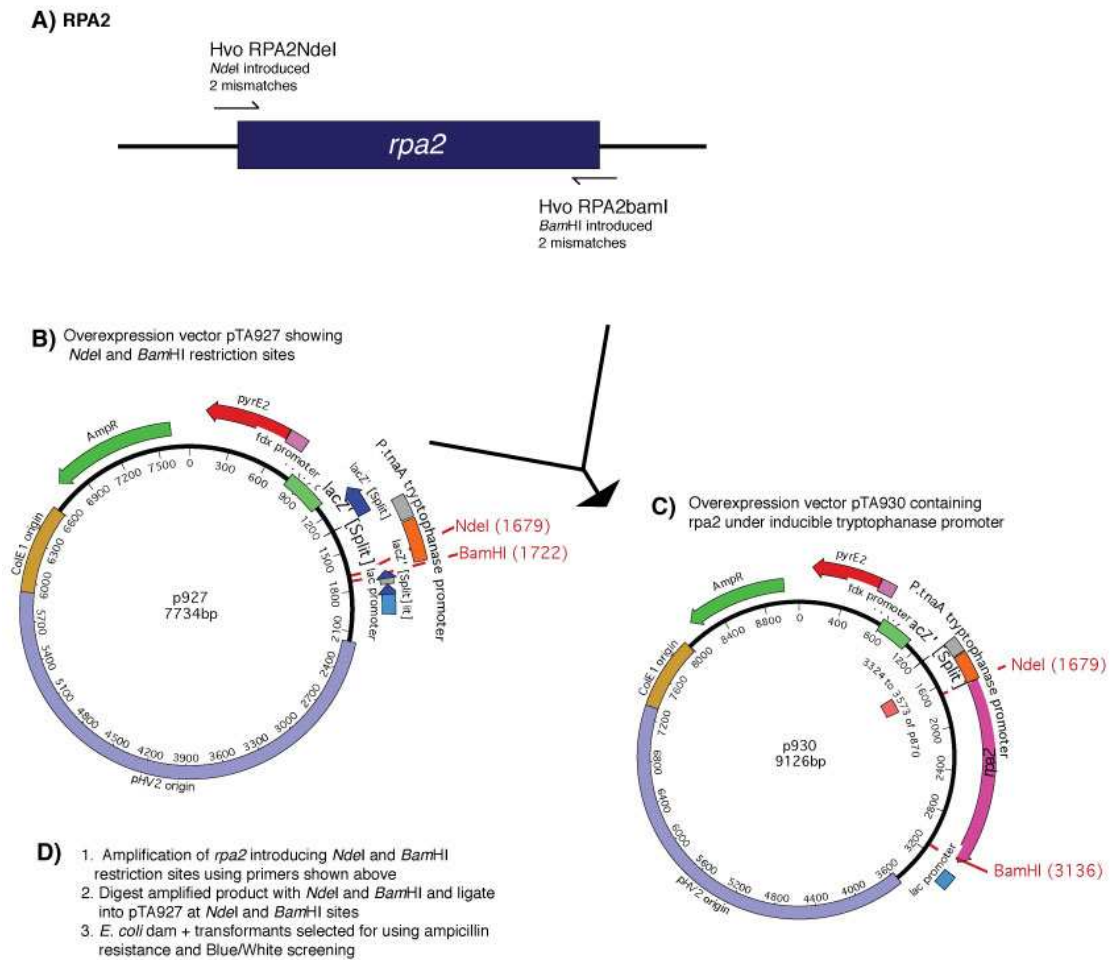
## 6.2 Overexpression of Rpa2

Overexpression of *E. coli* SSB has been shown to be detrimental to cell viability and growth (Moreau 1988). To examine if the effects of *rpa2* overexpression are similar those of *E. coli* SSB overexpression on cell viability and growth, *rpa2* was inserted into the overexpression plasmid pTA927 that contains a tryptophan-inducible tryptophanase promoter (Allers, Barak *et al.* 2010). The overexpression plasmid containing *rpa2*, pTA930, was constructed as shown in Figure 6.5.

Primer	Sequence 5'-3'	Restriction site inserted
Hvo RPA2NdeI	AGGTGACGCCat <b>ATG</b> GGCGTCAT	<i>NdeI</i>
Hvo RPA2bamI	GCGGTgg <b>ATC</b> CGTCCTACTGC	<i>BamHI</i>

**Table 6.1 Primers for the construction of *rpa2* overexpression plasmid**

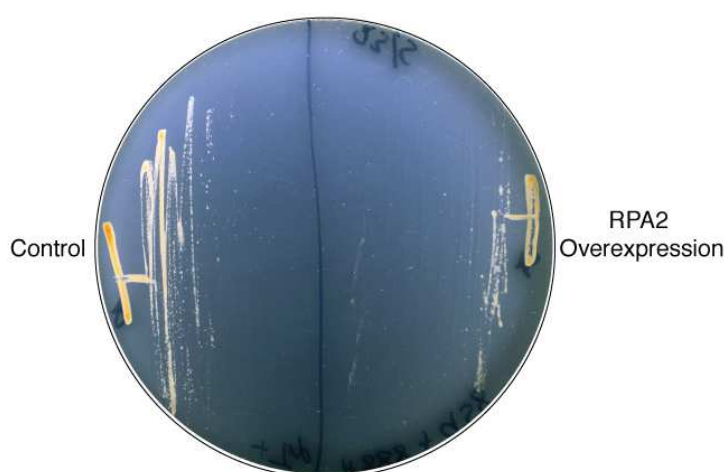
The sequence for each primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined). Start codon of *rpa2* is shown in bold.



**Figure 6.5 Construction of *rpa2* overexpression plasmid**

A) Binding sites of the primers are shown B) the restriction sites in the vector pTA927 used to insert *rpa2* C) the resulting *rpa2* overexpression construct pTA930. D) A basic sequence of construction.

After transformation into *E. coli* dam<sup>-</sup> the plasmid pTA938 was transformed into H888 (*trpA*<sup>+</sup>, allowing the strain to be grown in the absence of the inducer tryptophan) generating the strain H912. When H888 and H912 were streaked out side-by-side on Hv-Ca +trp, H912 shows a significant growth defect as shown in Figure 6.6. A similar growth defect was seen with overexpression of *E. coli* SSB (Moreau 1988), due to the binding of SSB to ssDNA and inhibition of DNA replication and repair.



**Figure 6.6 Overexpression of *rpa2***

Growth defect of the RPA2 overexpression strain H912 compared to that of the control (H888). Streaked out on Hv-Ca +trp.

## 6.3 Generation of *rpa2* deletion mutant

### 6.3.1 Construction of *rpa2* deletion mutant

#### Design of primers

Upstream and downstream regions of *rpa2* were amplified from the genomic *rpa2* clone pTA898 by touchdown PCR using the primers shown in Table 4.2.1, introducing external *KpnI* and *XbaI* restriction sites and an internal *BamHI* restriction site.

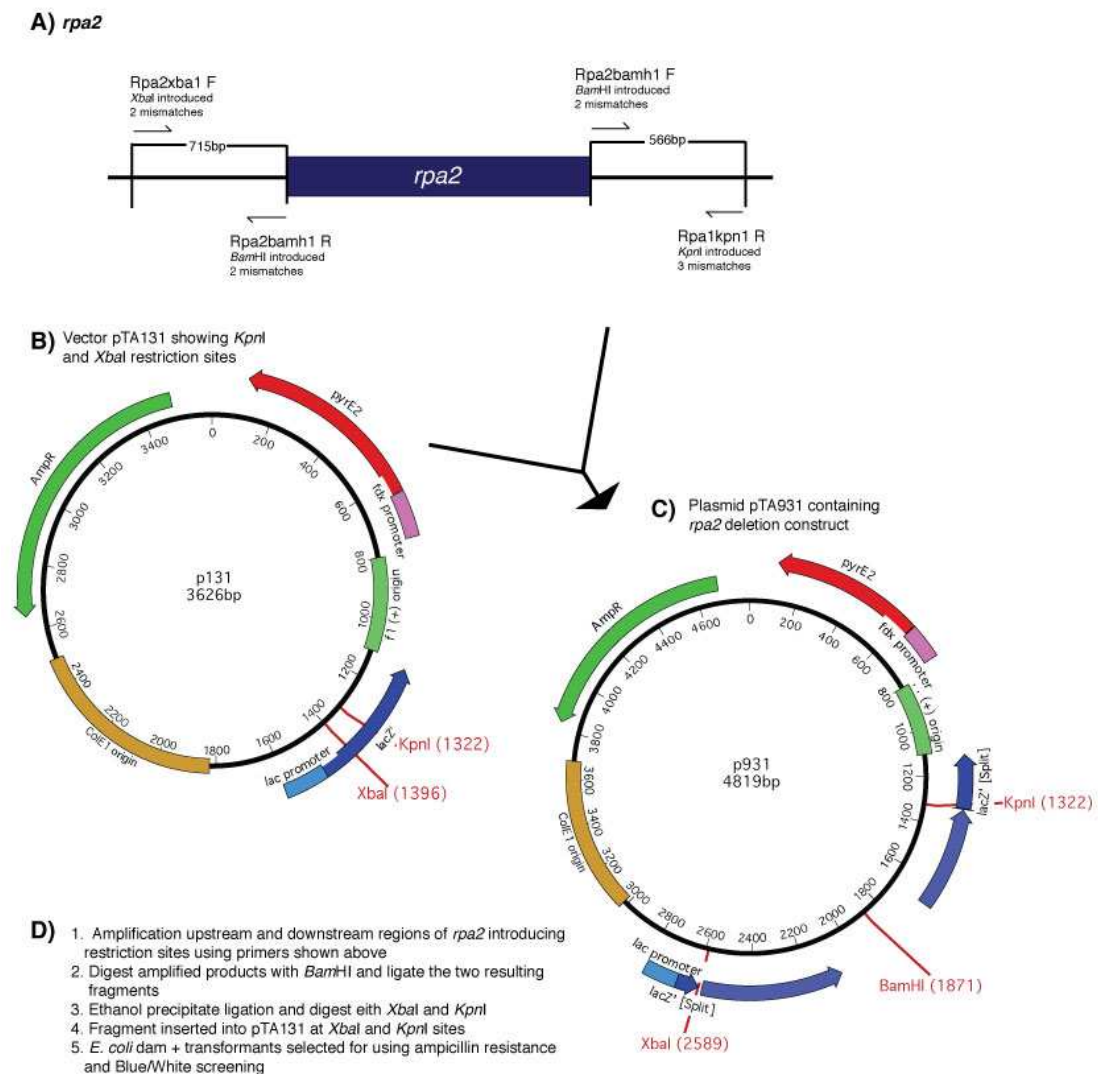
Primer	Sequence (5'-3')	Site inserted
Rpa2bamh1 F	CAGTAGGACGGATccACCGCCGGTC	<i>BamHI</i>
Rpa2kpn1 R	GCGTGGTaccCGAGTTTGACGGCGGC	<i>KpnI</i>
Rpa2xba1 F	CGGGtCTAGaTTCGCACGTGAATCG	<i>XbaI</i>
Rpa2bamh1 R	GCCTG <b>AT</b> GGCGggATCCGGGA	<i>BamHI</i>

**Table 6.2 Design of deletion primers for *rpa2***

The sequence for each deletion primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of the upstream (US) and downstream (DS) regions. Start codon of the *rpa2* shown in bold.

## Construction of *rpa2* deletion plasmid

The PCR products of the upstream and downstream regions were gel purified, digested with *Bam*HI and ligated together. The upstream/downstream ligation was ligated into the plasmid pTA131 using the *Kpn*I and *Xba*I sites, and was transformed into *E. coli* XL-1 strain, as shown in Figure 6.7. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA931 construct.

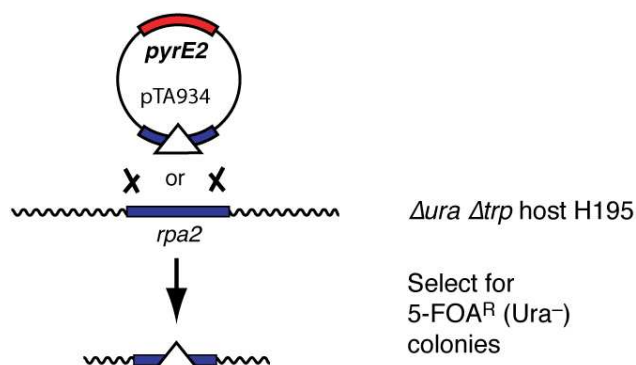


**Figure 6.7 Construction of the *rpa2* deletion**

A) Binding sites of the primers are shown B) the restriction site in the vector used to insert the upstream and downstream ligation C) the resulting deletion construct pTA931. D) A basic sequence of construction.

### Pop-in/ pop-out of the deletion construct

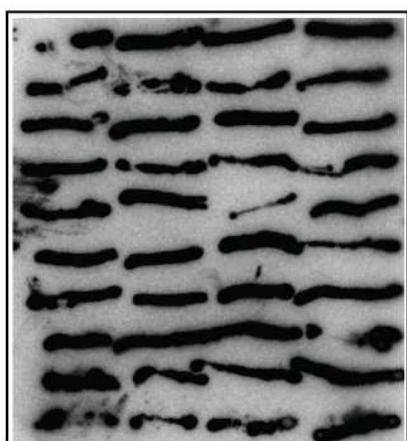
Transformation into *E. coli dam*<sup>-</sup> generated the plasmid pTA934, which was transformed into *H. volcanii* H195 generating the pop-in strain H891, to attempt to generate a genomic deletion of *rpa2* as shown in Figure 6.8.



**Figure 6.8 Schematic diagram of *rpa2* deletion**

The deletion plasmid pTA934 was constructed as described in Figure 6.7. Integration onto the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) should result in a  $\Delta rpa2$  strain.

In order to avoid merodiploid *H. volcanii* strains, potential deletions were screened by colony hybridisation, using a radioactive <sup>32</sup>P labelled *rpa2* probe. However all colonies hybridised with the probe as shown in Figure 6.9, most likely all the pop-outs had reverted to wild-type, or they could potentially have been merodiploid for native *rpa2* and  $\Delta rpa2$ .



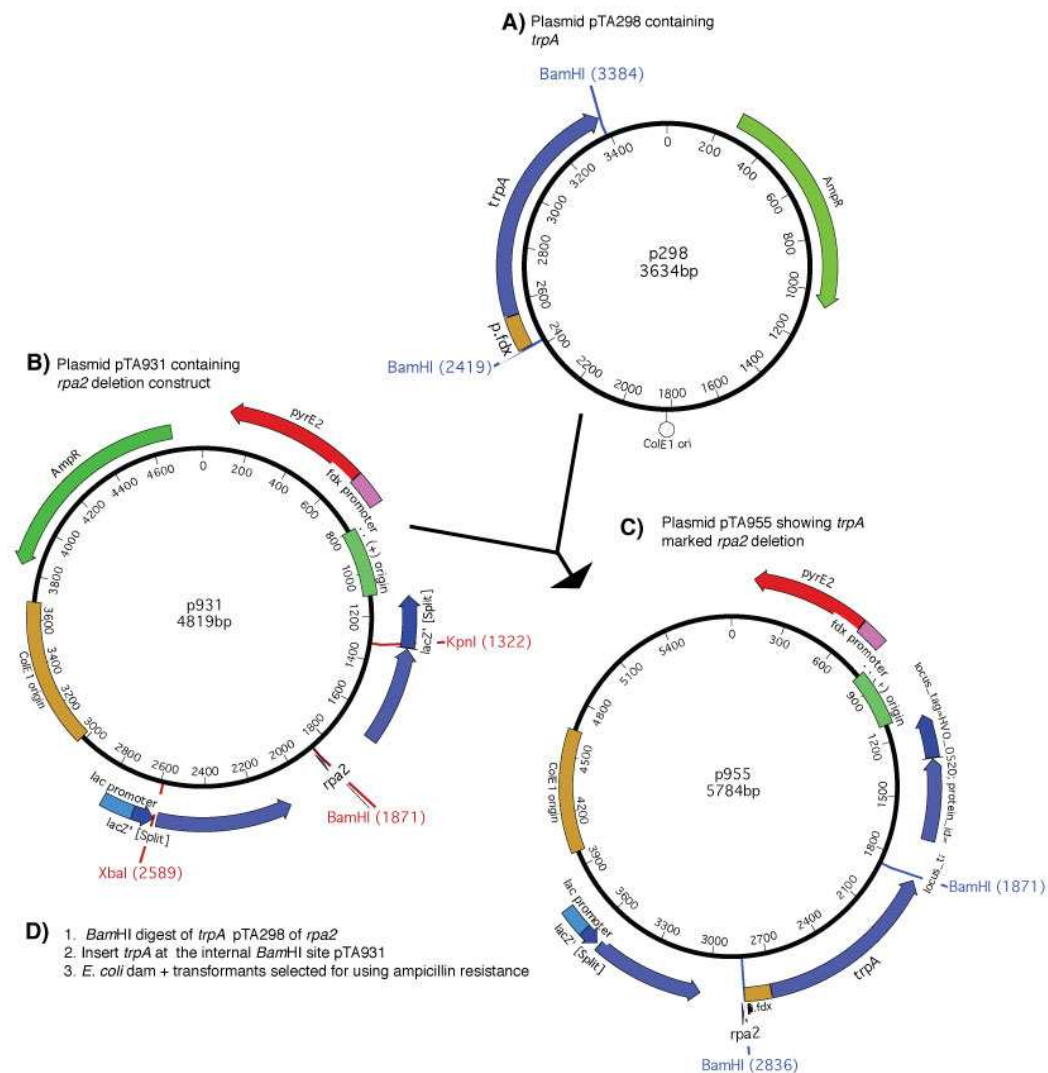
**Figure 6.9 Colony lift and hybridisation of H891 popouts**

Pop-outs (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv-YPC +thy. All pop-outs hybridised with the *rpa2* probe.

### 6.3.2 Construction of *trpA* marked *rpa2* deletion mutant

#### Construction of *rpa2* deletion plasmid

The *trpA* marker was digested from pTA298 using *Bam*HI and inserted into pTA931 at the internal *Bam*HI site. The resulting plasmid was then transformed into *E. coli* XL-1 strain. Restriction digest of plasmid extractions from transformants was used to screen for the correct construct pTA955, as shown in 6.10.



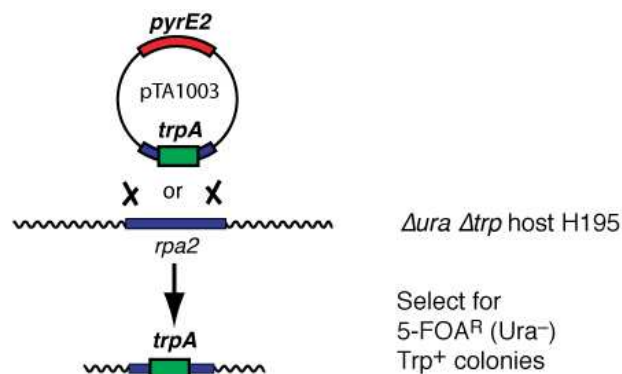
**Figure 6.10 Construction of the *trpA*-marked *rpa2* deletion**

A) *Bam*HI digest of *trpA* from pTA298 B) pTA931 showing internal *Bam*HI site C) *trpA* marked *rpa2* deletion construct pTA955 D) Basic sequence of construction.



### Pop-in/ pop-out of the deletion construct

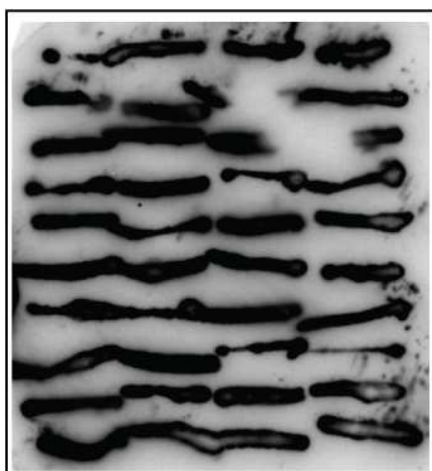
Following transformation into the *E. coli dam*<sup>-</sup> strain, the plasmid pTA1003 was transformed into *H. volcanii* H195, generating the pop-in H1025. This was used in an attempt to generate a genomic deletion of *rpa2* using the counter-selective pop-in/pop-out method, with additional selection for tryptophan as shown in Figure 6.11. This should bias the homologous recombination events to generate a genomic deletion of *rpa2* rather than reverting to the wild-type.



**Figure 6.11 Schematic diagram of *rpa2* deletion with selection for the *trpA* marker**

The deletion plasmid pTA1003 was constructed as described in Figure 6.10. Integration onto the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) with selection for Trp<sup>+</sup> should result in a  $\Delta rpa2::trpA^+$  strain.

Potential deletions were screened by colony hybridisation, using a <sup>32</sup>P labelled *rpa2* probe. However all pop-outs hybridised to the probe as shown in Figure 6.12. Therefore all the pop-outs are merodiploid for native *rpa2* and  $\Delta rpa2::trpA^+$ , suggesting that *rpa2* is essential for cell survival and growth.



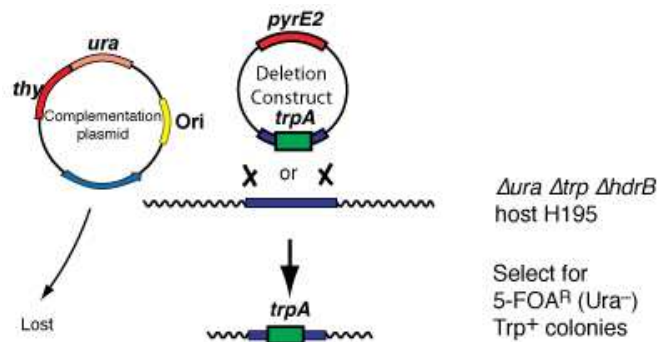
**Figure 6.12. Colony lift and hybridisation of H1025 popouts**

Pop-outs (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv-Ca +thy +ura, to maintain selection for *trpA* marked deletion construct. All pop-outs hybridised with the *rpa2* probe.



### 6.3.3 Construction of *rpa2* complementation plasmid

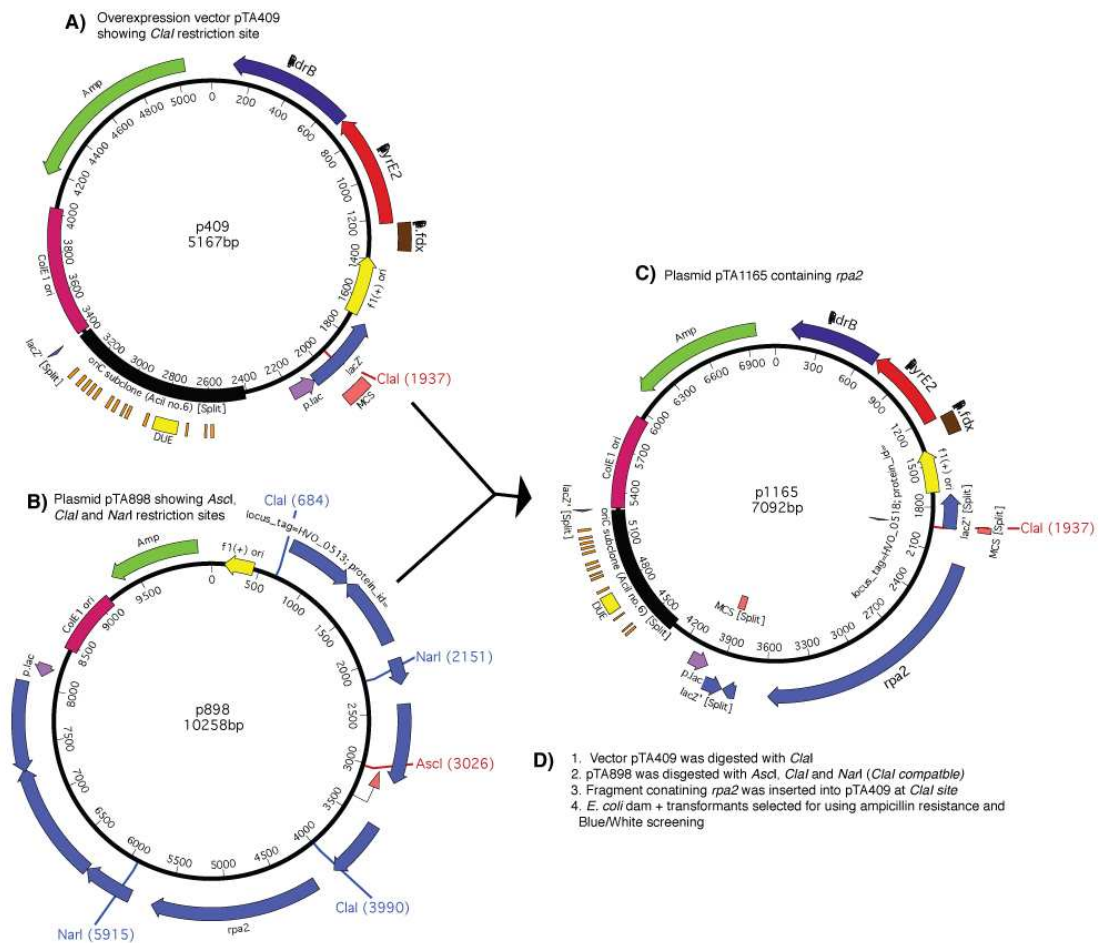
To maintain cell viability during the counter-selective pop-in/pop-out process, a plasmid for the expression *in trans* of Rpa2 was constructed. This may favour homologous recombination to generate a genomic deletion as opposed to reversion to the wild-type, since *rpa2* levels will be maintained by pTA1165. This technique, shown in Figure 6.13, had previously been used to generate a *radA* deletion in a  $\Delta mre11 \Delta rad50$  background (Delmas, Shunburne *et al.* 2009).



**Figure 6.13 Schematic of pop-in/pop-out with complementary *in trans* expression**

Figure shows the use of complementary *in trans* expression during the pop-out process, the episomal plasmid is positively selected for using  $\text{Thy}^+$  and is lost by removal of  $\text{Thy}^+$  selection and 5-FOA<sup>R</sup> selection. Adapted from (Leigh, Albers *et al.* 2011).

The *rpa2* gene was inserted into a shuttle vector pTA409, which features *hdrB* and *pyrE2* markers, generating the plasmid pTA1165. The *hdrB* marker allows selection for the episomal plasmid in a *trpA*-marked pop-in strain, while the *pyrE2* marker allows counter-selection for loss of the episomal plasmid by 5-FOA sensitivity after the pop-out recombination event has occurred. The restriction enzymes *NarI* and *ClaI* were used to digest *rpa2* from the genomic *rpa2* clone, pTA898, this fragment was inserted at a compatible *ClaI* site in pTA409, generating the plasmid pTA1165 as shown in Figure 6.14.

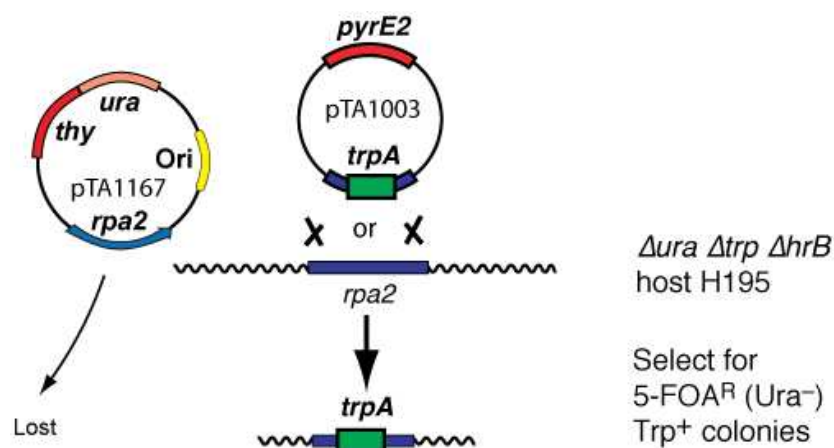


**Figure 6.14 Construction of the *rpa2* complementation plasmid**

A) Vector pTA409 showing *Clal* site B) pTA898 showing *Clal*, *NarI* (*Clal* compatible) and *AscI* site, used to cut unwanted fragments into smaller sized fragments to distinguish them from the fragment containing *rpa2*. C) *In trans* expression plasmid pTA1165 containing *rpa2*, constructed by ligation of *NarI* site to *Clal* site.

### Pop-in/ pop-out of the deletion construct

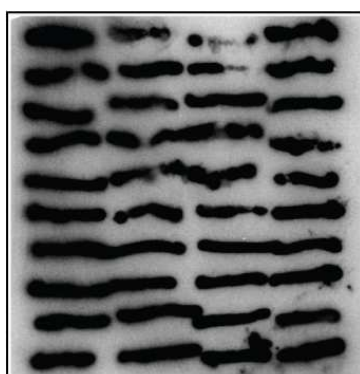
Following transformation into the *E. coli* dam<sup>-</sup> the episomal complementation plasmid pTA1167 was transformed into *H. volcanii* pop-in strain H1025, generating the strain H1193, which was then used to attempt to generate a genomic deletion of *rpa2* as shown in Figure 6.15.



**Figure 6.15 Schematic diagram of *rpa2* deletion with complementation *in trans***

The episomal plasmid pTA1167, constructed as described in Figure 6.14, expresses *rpa2* *in trans* during the pop-out process and is lost by removal of Thy<sup>+</sup> selection and 5-FOA<sup>R</sup> selection.

Potential deletions were screened by colony hybridisation, using a <sup>32</sup>P labelled *rpa2* probe. However all pop-outs hybridised to the probe, indicating that the complementation *in trans* of *rpa2* was unsuccessful and consequently *rpa2* is still present.



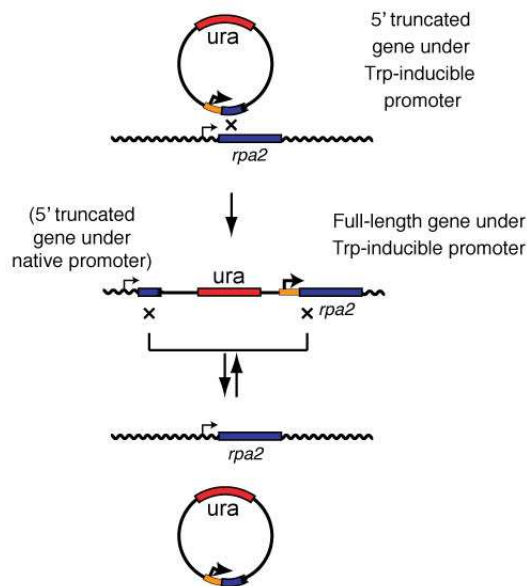
**Figure 6.16 Colony lift and hybridisation of H1193 popouts**

Popouts (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv-YPG. All pop-out candidates hybridised with the *rpa2* probe.

## 6.4 Integration of *rpa2* at an ectopic locus

The failure of the three methods described above to generate an *rpa2* deletion mutant suggests that *rpa2* is essential. If *rpa2* is essential then it is likely that a total loss of *rpa2* expression would result in cell death, and consequently the exact role of Rpa2 could not be investigated. Alternatively, a two-step conditional knockdown method was used, involving the integration of *rpa2* (under a tryptophan-inducible promoter) at an ectopic location, followed by the attempted deletion of native *rpa2* (using a *hdrB*

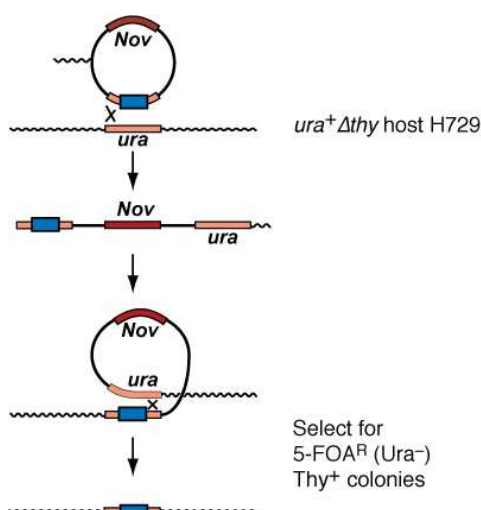
selection). This method has the advantage of stability of over the one-step procedure used previously by Large *et al* (2007), which involves the replacement of the native promoter with a tryptophan-inducible promoter (Large, Stämme *et al.* 2007). This can revert back to the wild-type, where the gene is under its native promoter, by homologous recombination as shown in Figure 6.17 (Thorsten Allers, unpublished).



**Figure 6.17 Instability of native promoter replacement**

Figure shows the integration of a tryptophan-inducible promoter and the 5' portion of *rpa2* at the *rpa2* locus, replacing the native *rpa2* promoter and 5' portion of the gene. However the native promoter and the 5' portion are still present, therefore homologous recombination between the two 5' portions can restore the native *rpa2* promoter, excising the plasmid from the chromosome.

The two-step method used here provides positive selection at the first step for integration at the ectopic location (*pyrE2*), by using novobiocin resistance (Figure 6.18). At the second stage, where the native *pyrE2* is deleted, 5FOA resistance is used for selection. This method will introduce *rpa2* under an inducible promoter at the *pyrE2* locus, generating a conditional knockdown and thus allowing the role of *rpa2* in DNA replication and repair to be examined.

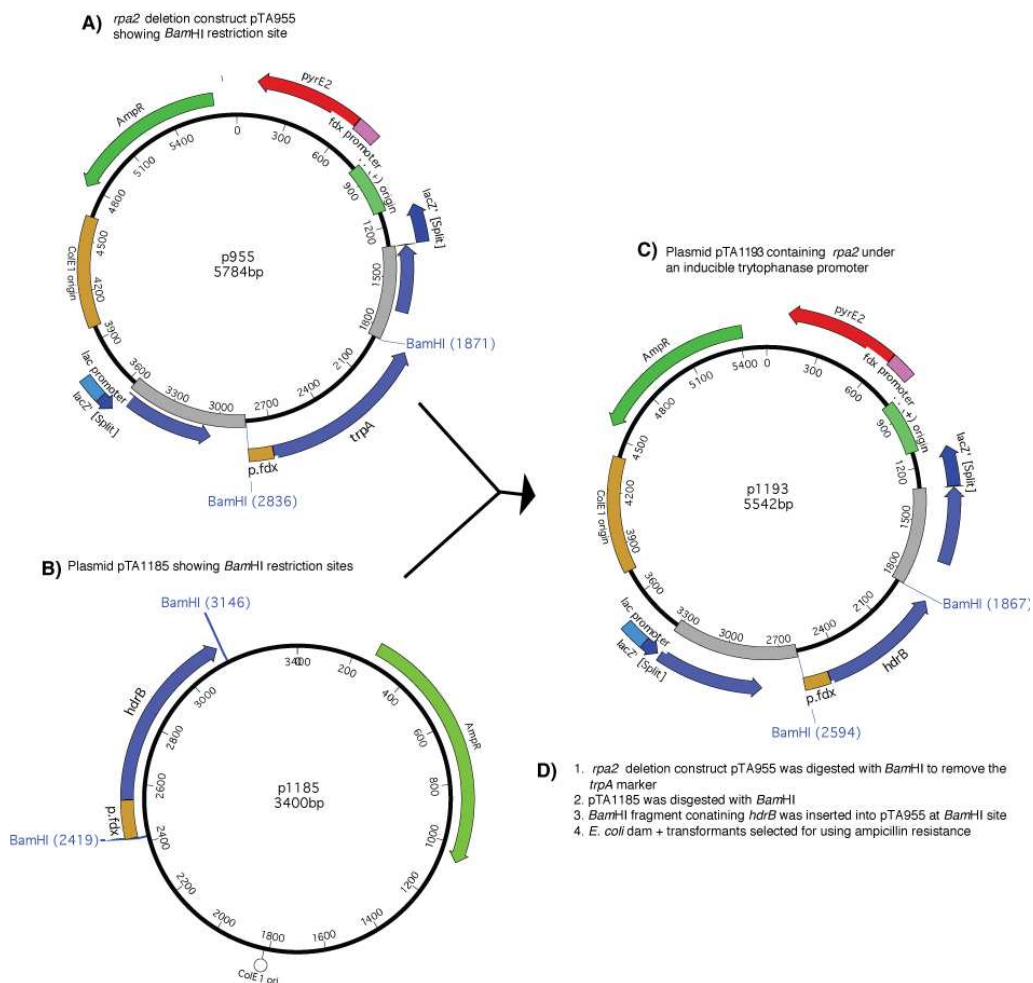


**Figure 6.18 Inserting a gene at an ectopic locus**

The plasmid integrates at the ectopic locus via homologous recombination, the resulting pop-in is selected for using novobiocin resistance (Nov<sup>R</sup>). During the pop-out process, loss of the plasmid and replacement of *pyrE2* is selected for using 5-FOA<sup>R</sup>. Adapted from (Leigh, Albers *et al.* 2011).

#### **6.4.1 Construction of the $\Delta rpa2$ *hdrB* marked plasmid**

The *trpA* marker cannot be used in combination with the *rpa2* at an ectopic locus if the latter is under control of the tryptophanase promoter. The strain must be *trp*<sup>+</sup> in order to be viable in the absence of tryptophan, since it will be necessary to remove tryptophan in order to repress the *rpa2* gene and investigate the function of RPA2. Furthermore, marking the *rpa2* deletion with *trpA* under the constitutive ferredoxin promoter would prevent this form of gene regulation. Instead, an *hdrB* marked *rpa2* deletion construct was made. The *hdrB* marker was excised from pTA1185 using *Bam*HI and inserted into the internal *Bam*HI site in pTA955. After transformation into *E. coli* XL-1, restriction digest of plasmid extractions from transformants was used to screen for the correct pTA1193 construct, shown in Figure 6.19.



**Figure 6.19 Construction of the *hdrB* marked *rpa2* deletion**

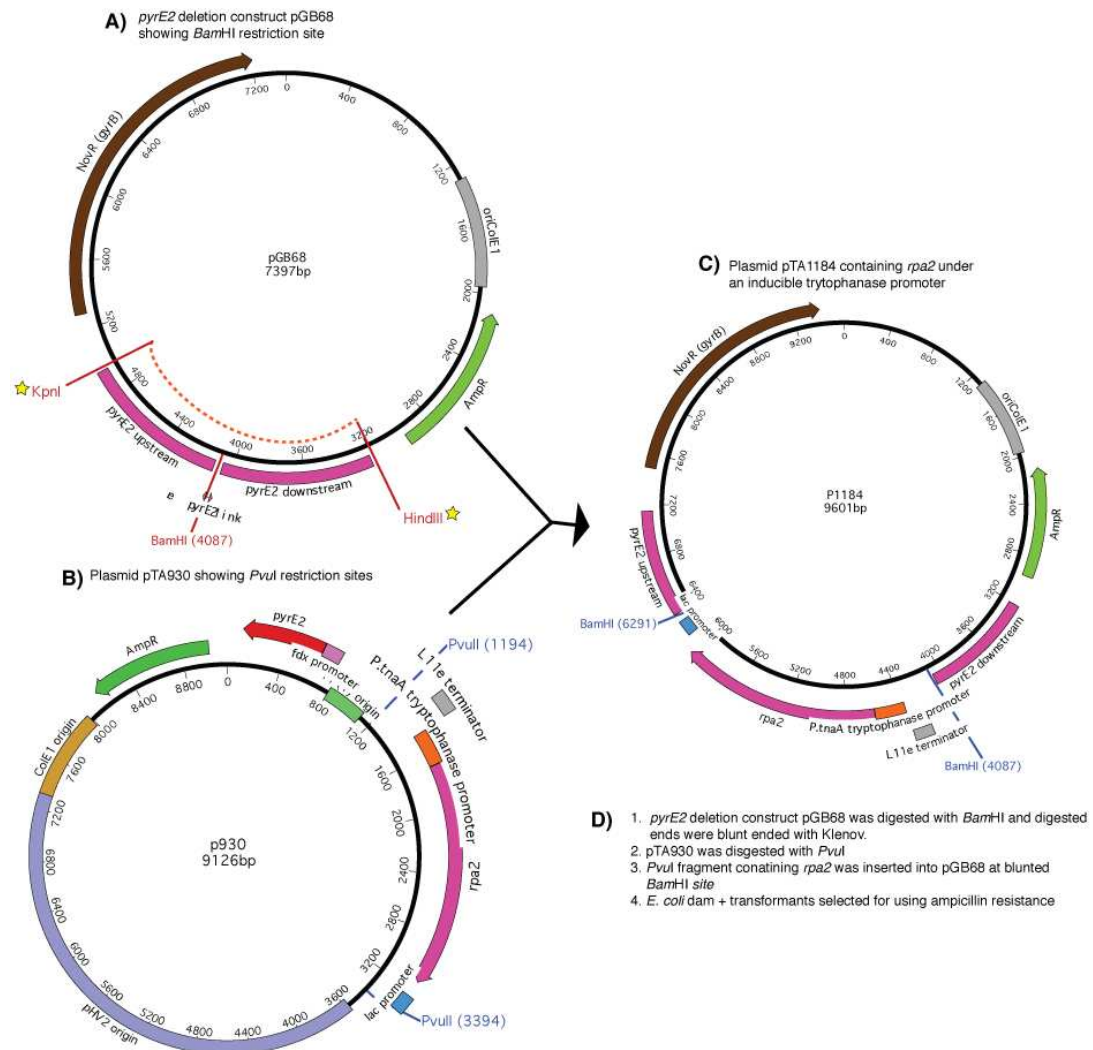
Deletion construct of *rpa2*, pTA955 showing *Bam*HI site. B) pTA1185 containing *hdrB* marker, showing *Bam*HI sites C) *rpa2* deletion construct with insertion of *hdrB* D) A basic sequence of construction.

#### 6.4.2 Insertion of *rpa2* at an ectopic locus

The *pyrE2* gene was chosen as the ectopic locus for *rpa2* insertion, since it would allow selection for 5FOA resistance at the pop-out stage, as shown in Figure 6.19. To generate a construct that would replace *pyrE2* with inducible *rpa2*, a fragment containing the *rpa2* gene and the tryptophanase promoter was excised from pTA930 (shown in Figure 6.5, details of the primers used are shown in Table 6.1) and inserted in the *pyrE2* deletion construct pGB68 (Bitan-Banin, Ortenberg *et al.* 2003).

The restriction enzyme *Pvu*II was used to excise *rpa2* from pTA930 (see Figure 6.5), as shown in figure 6.21. *Pvu*II is a blunt-end restriction enzyme, therefore before the *Pvu*II fragment could be inserted at an internal *Bam*HI site in the *pyrE2* deletion

construct pGB68 (Bitan-Banin, Ortenberg *et al.* 2003), the sticky *Bam*HI ends needed to be ‘filled-in’ by the Klenow polymerase. After ligation of the *Pvu*II fragment containing inducible *rpa2* with pGB68, the plasmid was transformed into *E. coli* XL-1, followed by restriction digest of plasmid extractions from transformants to screen for the correct pTA1184 construct, as shown in Figure 6.20.



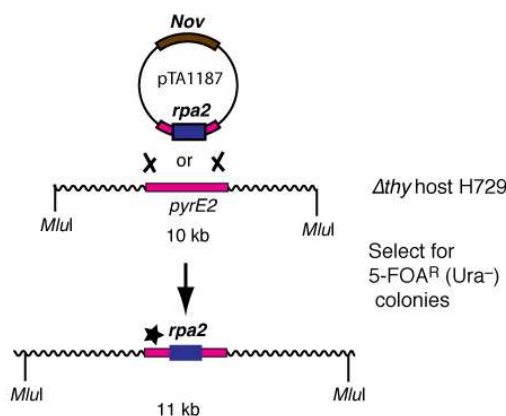
**Figure 6.20 Construction of the *pyrE2* deletion marked with inducible *rpa2***

A) The deletion construct of *pyrE2*, pGB68 showing *Bam*HI site. B) Overexpression plasmid pTA930, shown in Figure 6.5, showing *Pvu*II sites used to excise *rpa2* under the tryptophanase promoter. C) *pyrE2* deletion construct with insertion of *rpa2* under the tryptophanase promoter D) A basic sequence of construction. Orange dotted line and yellow stars denotes region that was used as a <sup>32</sup>P labelled probe in Southern hybridisation shown in Figure 6.24.



### Pop-in/ pop-out of the $\Delta pyrE2::rpa2+$ construct

Following transformation into the *E. coli dam*<sup>-</sup>, the plasmid pTA1187 was transformed into *H. volcanii* H729 (*pyrE2*<sup>+</sup> $\Delta hdrB$ ) generating the pop-in strain H1234, which was used to attempt to generate a genomic deletion of *pyrE2* and integration of *rpa2* (under control of the tryptophanase promoter) at the *pyrE2* locus, as shown in Figure 6.21.

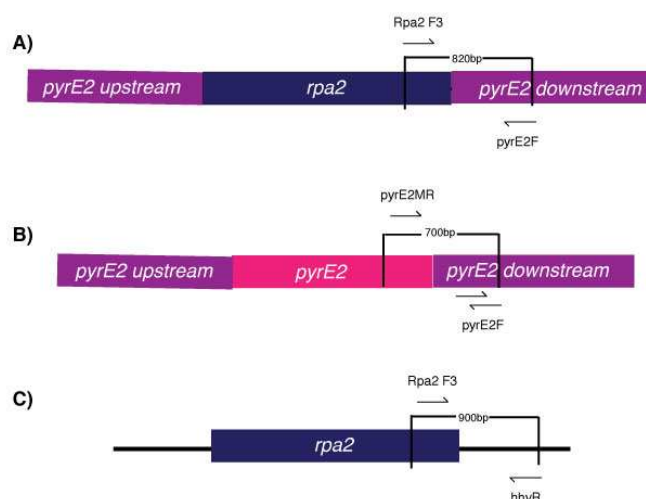


**Figure 6.21 Schematic diagram of *pyrE2* replacement with inducible *rpa2***

The gene replacement plasmid pTA1187 was constructed as described in Figure 6.20. Restriction digest with *MluI* (see text for details) and the resulting fragments expected are shown. Integration into the genome by homologous recombination (pop-in) using novobiocin selection, followed by loss of the plasmid by homologous recombination (pop-out) will result in the  $\Delta pyrE2$  deletion or a reversion back to wild-type. The region that the <sup>32</sup>P labelled probe hybridises to is denoted by a star.

It was not possible to screen by colony hybridisation for insertion of *rpa2* at the *pyrE2* locus, since the native *rpa2* is still present. To screen for insertion of the inducible *rpa2* (under control of the tryptophanase promoter) at the *pyrE2* locus, genomic PCRs were performed using an external *pyrE2* primer and an internal *rpa2* primer, as shown in Figure 6.22. A positive control was used for each primer on wild-type *H. volcanii* H678 DNA (*pyrE2*<sup>+</sup>, *rpa2*<sup>+</sup>), as shown in Figure 6.22.

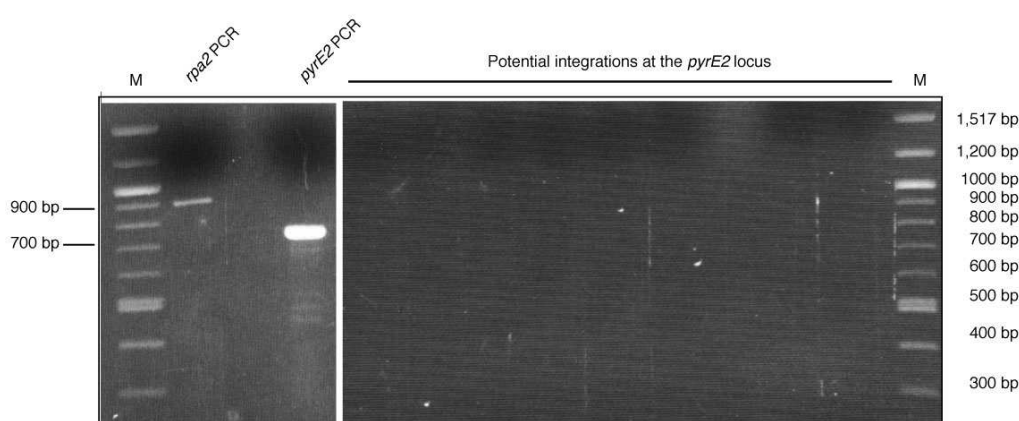




**Figure 6.22 Design of diagnostic PCRs to check *rpa2* integration at *pyrE2* locus**

A) Diagnostic PCR to check for inducible *rpa2* integration at the *pyrE2* locus. B) Control PCR to check *pyrE2F* primer binding. C) Control PCR to check *Rpa2 F3* primer binding.

Despite the positive controls generating a PCR product of the expected size as shown in Figure 6.23, all of the potential pop-outs failed to generate a PCR product, indicating that they had all reverted back to wild-type. It should be noted that the novobiocin marker was used, since it allows positive selection for the pop-in of inducible *rpa2* at the *pyrE2* locus (Figure 6.21). However it can have a toxic effect on the cell, by inhibiting the B subunit of the DNA gyrase (Gellert, O'Dea *et al.* 1976). It is possible that *pyrE2* may have been mutated as a consequence of novobiocin exposure, resulting in 5FOA resistance and an unstable pop-in, thus making it likely that plasmid pTA1187 would be excised from the chromosome.

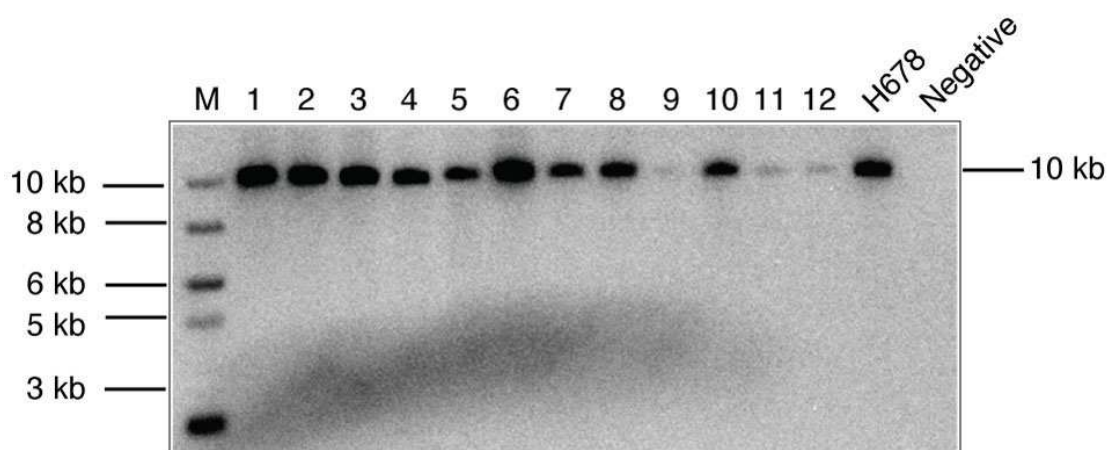


**Figure 6.23 Control PCRs for *rpa2* integration at the *pyrE2* locus.**

M shows 100 bp marker, predicted sizes for *rpa2*, *pyrE2* and *rpa2-pyrE2* PCR are shown in Figure 6.22.

To overcome this problem and avoid mutation of *pyrE2*, the novobiocin selection for pop-ins was not used. Instead the transformation of H729 with pTA1187 was repeated, but instead of propagating the pop-in transformants as usual (on Hv-Ca +thy +Nov plates, followed by non-selective growth for the pop-out process), they were immediately plated out on Hv-Ca +5FOA +thy. This should result in a pop-in/pop-out in one step by selecting for *pyrE2* deletion. This method of direct gene replacement has been used successfully in other laboratories (Moshe Mevarech, Tel Aviv University, personal communication).

The genomic sequence of *H. volcanii* was used to determine an appropriate restriction digest to screen the potential deletion mutants, this was then used in a Southern hybridisation to confirm genomic deletion of *pyrE2* and the insertion of inducible *rpa2*. A <sup>32</sup>P labelled probe of the upstream and downstream region of *pyrE2* was used in the hybridisation, shown in Figure 6.24. However all of the potential pop-outs were positive for native *pyrE2*, as shown in Figure 6.24.



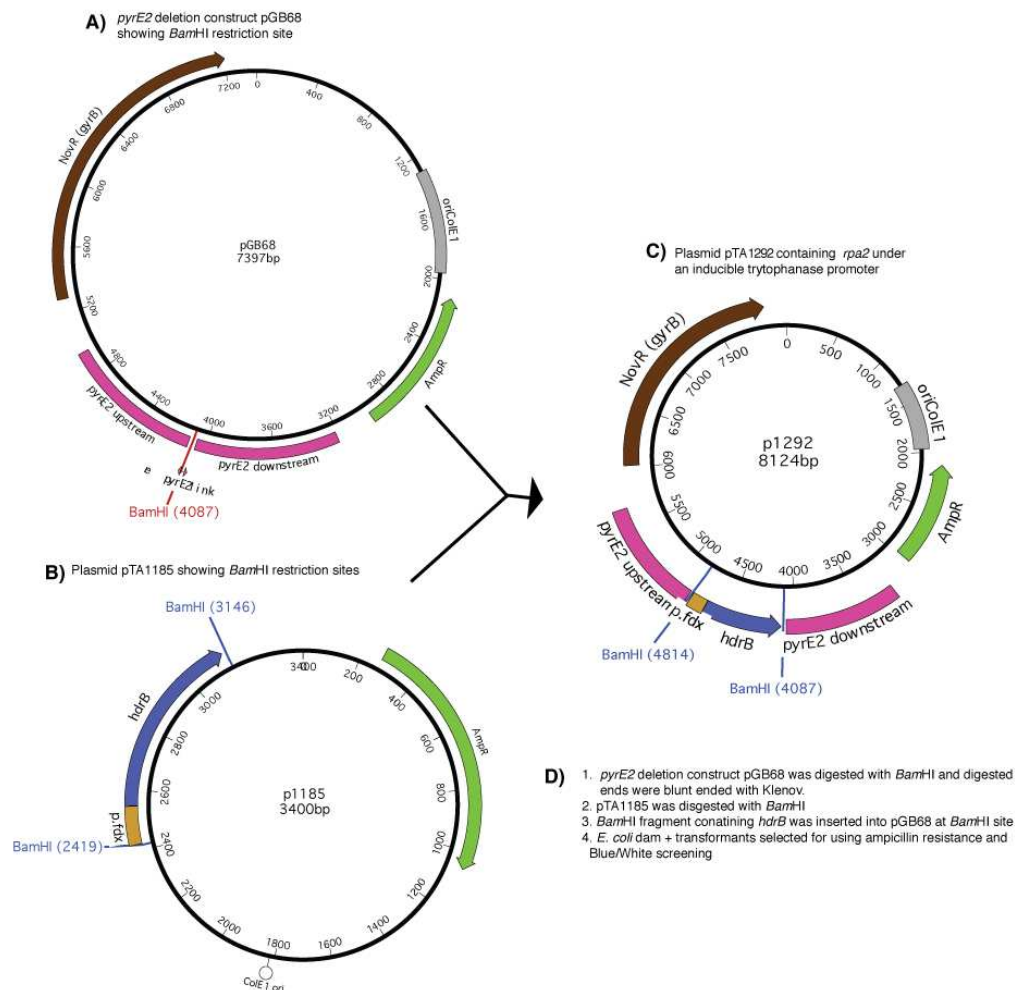
**Figure 6.24 Southern blot analysis of potential H1234 pop-outs**

Lane M shows size marker lane. Genomic DNA was digested with *MluI* and probed with the flanking regions of the deletion construct pGB68 digested with *EcoRI/XbaI*. Genomic DNA of H678 was used as a wild-type control. See Figure 6.21 for further restriction digest details.

## 6.5 Insertion of *hdrB* at the *pyrE2* locus

As a positive control for integration of a gene at an ectopic locus, the plasmid pTA1292 was constructed as shown in Figure 6.25, for the insertion of the thymidine biosynthesis gene *hdrB* at the *pyrE2* locus.

### 6.5.1 Construction of *pyrE2* deletion construct marked with *hdrB*

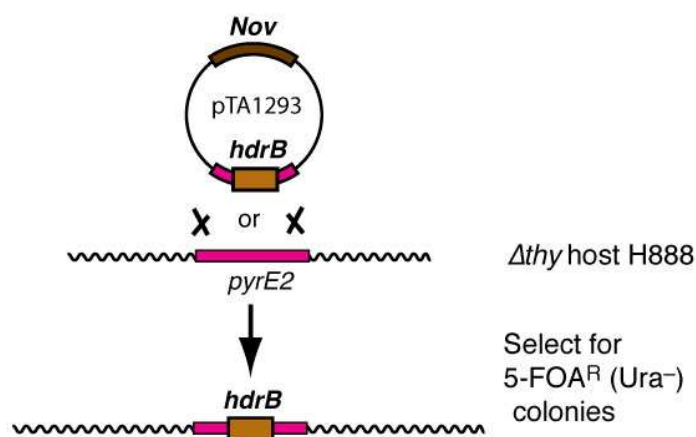


**Figure 6.25 Construction of the *pyrE2* deletion marked with *hdrB***

A) The *pyrE2* deletion construct pGB68, showing *Bam*HI site. B) Plasmid pTA1185 showing *Bam*HI sites used to excise *hdrB* under control of the ferredoxin promoter. C) *pyrE2* deletion construct with insertion of *hdrB*. D) Basic sequence of construction.

### Pop-in/ pop-out of the deletion construct

Following transformation into the *E. coli dam*<sup>-</sup> the plasmid pTA1293 was transformed into *H. volcanii* H729 (*pyrE2*+ $\Delta$ *hdrB*) generating the pop-in strain H1493, which was used to generate a genomic deletion of *pyrE2* and integration of *hdrB* at the *pyrE2* locus in strain H1494, as shown in Figure 6.26.

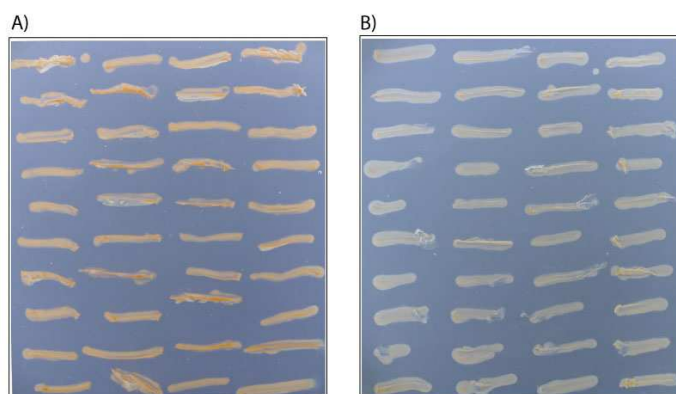


**Figure 6.26. Schematic diagram of *pyrE2* replacement with *hdrB***

The gene replacement plasmid pTA1293 was constructed as described in Figure 6.25. Integration into the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) will result in the  $\Delta$ *pyrE2* deletion marked with *hdrB*<sup>+</sup>, or a reversion back to wild-type.

### **6.5.2 Autotrophic conformation of *hdrB* insertion at the *pyrE2* locus**

Pop-outs were patched out on Hv-YPC +thy and Hv-YPC to test whether it is possible to insert *hdrB* at the *pyrE2* locus, as shown in Figure 6.27. All patches showed prototrophy for thymidine, indicating that the method of inserting a gene at an ectopic locus such as *pyrE2* is possible. However, given that it was unsuccessful with inducible *rpa2*, it is likely that positive selection for the insertion is required.

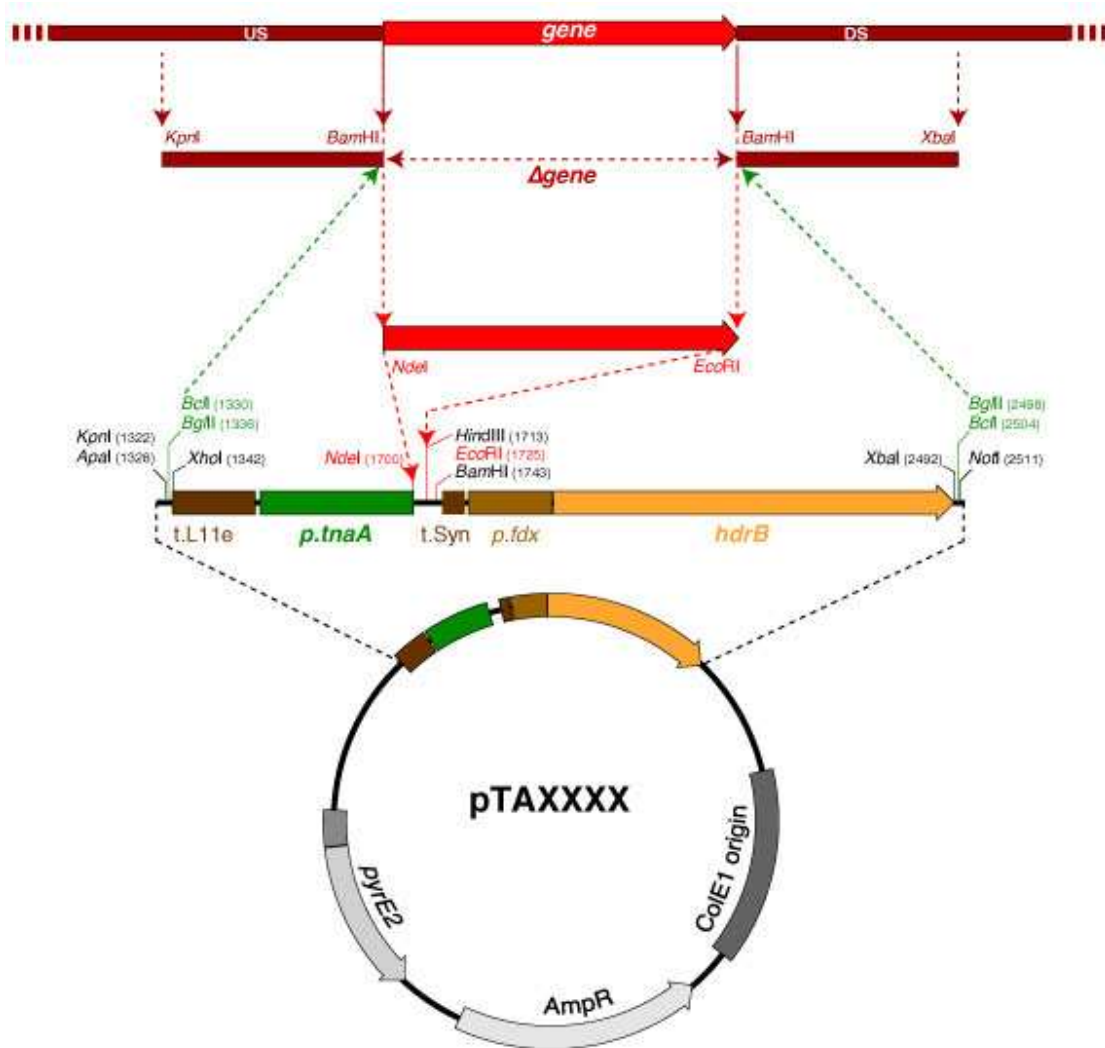


**Figure 6.27 H1494 patch plates on Hv-YPC +/- thy**

A) Shows pop-out colonies on Hv-YPC +thy. B) Shows the same pop-outs patched on Hv-YPC. All pop-outs show prototrophy for thymidine, confirming the *hdrB* marker had inserted at the *pyrE2* locus.

## 6.6 Discussion

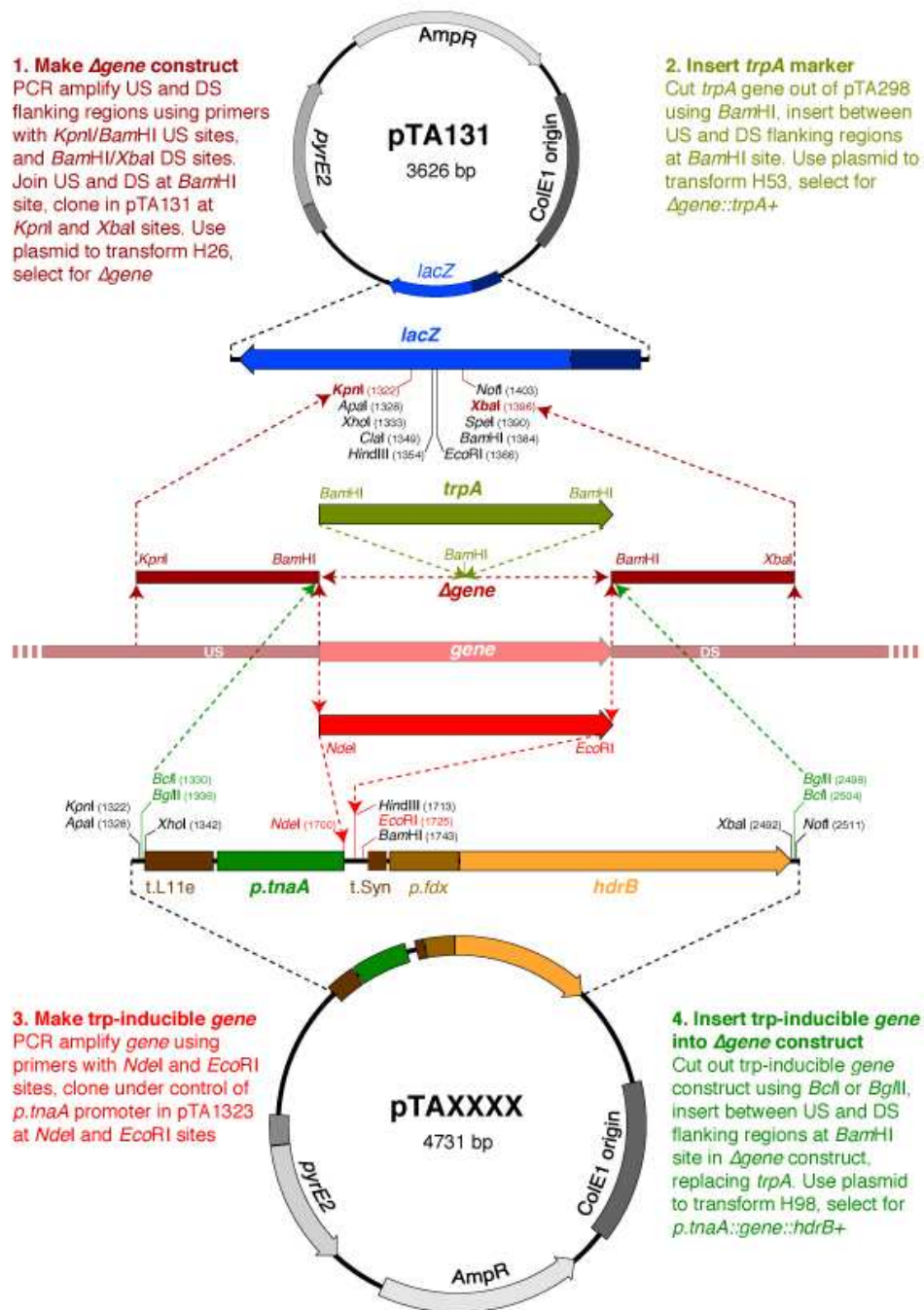
The failure to generate a genomic deletion of *rpa2* suggests that RPA2 is essential for cell survival, unlike RPA1 and RPA3. This indicates that RPA2 has a different cellular role to RPA1 and RPA3, therefore the three RPAs do not function as a heterotrimer as in eukaryotes. This assertion is also supported by the phylogenetic analysis showing RPA2 to be distantly related to RPA1 and RPA3. The failure to insert *rpa2* ectopically at the *pyrE2* locus (when *hdrB* could be inserted at this locus) suggests the expression of *rpa2* is tightly regulated; too little RPA2 in the cell would leave ssDNA unstable and susceptible to nuclease attack, whereas too much RPA2 would out-compete DNA replication and repair proteins in ssDNA-binding. Consequently, the presence of two copies of *rpa2* in the cell would be detrimental, which is supported by the growth defect exhibited by the RPA2 overexpression strain H912. Furthermore, the use of the selectable marker novobiocin to select for the pop-in stage of inducible *rpa2* at the *pyrE2* locus might have been problematic, since novobiocin inhibits DNA gyrase and could lead to an increased rate of mutation at *pyrE2*, generating spontaneous 5FOA resistance and thereby rendering the selection impossible. These problems could be overcome if *hdrB* positive selection were used for insertion of a tryptophan-inducible *rpa2* at the native *rpa2* locus. A schematic for the construction of such a plasmid is shown in Figures 6.28 and 6.29, however due to time constraints it was not possible to construct this plasmid.



**Figure 6.28 New concept for replacing native *rpa2* with an inducible *rpa2***

The schematic shows the gene of interest (*rpa2* in this case, light red) under control of the tryptophan inducible promoter (green) and linked to the selectable marker *hdrB* (orange). The inducible gene construct is flanked by the upstream and downstream regions of the native gene (dark red). These flanking regions allow integration of the construct at the native gene locus by homologous recombination (pop-in), which is selected for by uracil prototrophy, followed by the generation of a pop-out that is selected for by 5FOA resistance. Maintaining selection for thymidine prototrophy (*hdrB*<sup>+</sup>) should ensure that the inducible version of the gene is obtained.





**Figure 6.29 Summary of improved gene deletion/replacement concept**

Method for gene deletion (with optional positive selection for *trpA*) and/or conditional knockdown (with positive selection for *hdrB*). See Figure 6.28 for more details.

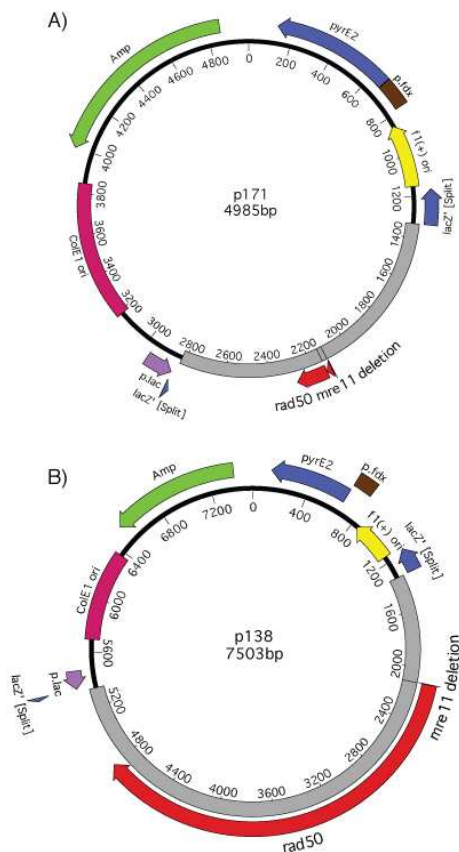


## Chapter 7: Multiple deletion mutants

### 7.1. $\Delta rpe \Delta mre11 \Delta rad50$ and $\Delta rpe \Delta mre11$ deletion mutants

#### 7.1.1. Generation of $\Delta rpe \Delta mre11 \Delta rad50$ and $\Delta rpe \Delta mre11$ deletion mutants

In eukaryotes, the Mre11-Rad50 complex is involved in the initial processing of damaged DNA ends during DNA repair, as discussed earlier. The Mre11 component possesses 3'-5' exonuclease and endonuclease activities and makes the initial incisions in the processing of the DNA ends at a DSB (Symington 2002). It is possible that the RPE might also act directly on damaged DNA, using its phosphoesterase activity to resect 'dirty' DNA ends. RPE features a calcineurin-like phosphoesterase domain, which could provide RPE with a nuclease activity that could function in the processing of DNA ends at a DNA break. To examine if RPE has a redundant function with Mre11 (and possibly Rad50) in the processing of damaged DNA ends, the double  $\Delta rpe \Delta mre11$  and the triple  $\Delta rpe \Delta mre11 \Delta rad50$  deletion mutants were made, using the constructs shown in Figure 7.1.

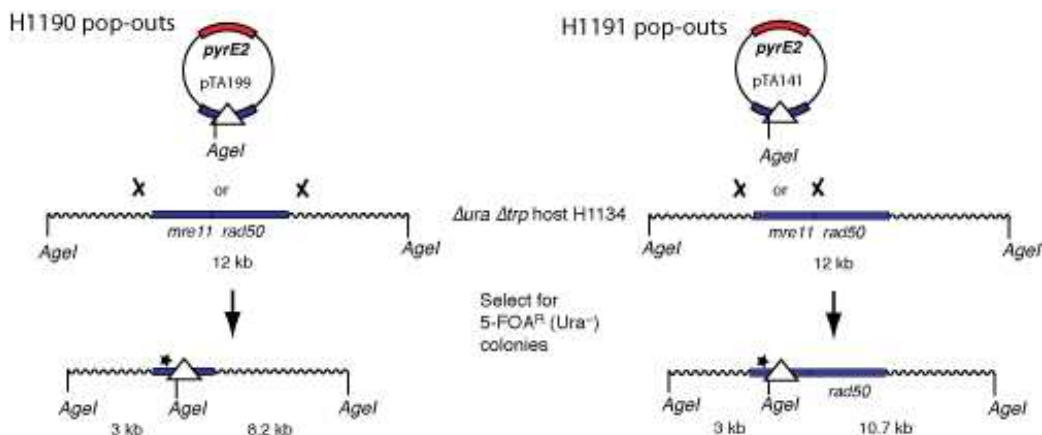


**Figure 7.1 *mre11-rad50* and *mre11* deletion constructs**

A) *mre11-rad50* deletion construct pTA171 (*dam*- plasmid pTA199). B) *mre11* deletion construct pTA138 (*dam*- plasmid pTA141) (Delmas, Shunburne *et al.* 2009).

### Pop-in/ pop-out of the deletion construct

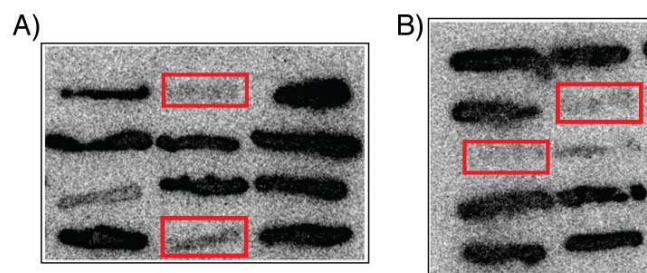
In *H. volcanii*, *mre11* is in an operon with *rad50*. The *mre11-rad50* deletion plasmid pTA199 and the *mre11* deletion plasmid pTA141 (Delmas, Shunburne *et al.* 2009), shown in Figure 7.1, were transformed into the  $\Delta rpe$  strain H1134, selecting for uracil prototrophy (*pyrE2*+) as shown in Figure 7.2, to generate the pop-in strains H1190 and H1191 respectively. The pop-in strains were then used to generate the deletion mutants  $\Delta rpe \Delta mre11 \Delta rad50$  (H1214) and  $\Delta rpe \Delta mre11$  (H1215) using the counter-selective pop-in/pop-out method discussed previously (Allers, Ngo *et al.* 2004).



**Figure 7.2. Schematic diagram of *mre11* and *mre11 rad50* deletion constructions**

The deletion plasmid pTA199 was transformed into H1134 generating the pop-in H1190, and the deletion plasmid pTA141 into H1134 generating the pop-in H1191. Restriction digest with *AgeI* and the resulting fragments expected that were used in the Southern blot analysis are shown. Integration into the genome by homologous recombination (pop-in) followed by loss of the plasmid by homologous recombination (pop-out) in the  $\Delta rpe \Delta mre11 \Delta rad50$  and  $\Delta rpe \Delta mre11$  deletions respectively. The region that the  $^{32}\text{P}$  probe hybridises to is denoted by a star.

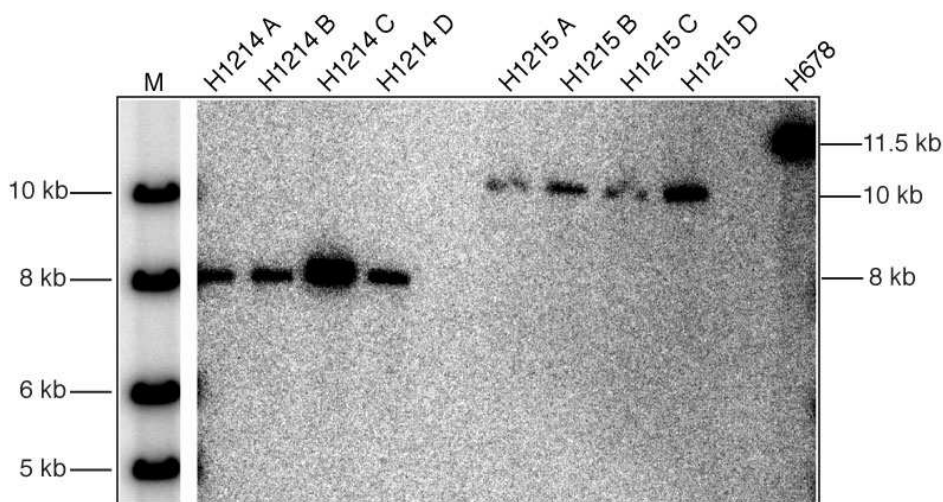
In order to avoid merodiploid *H. volcanii* strains, potential deletions were screened by colony hybridisation, using a radioactive  $^{32}\text{P}$  labelled *mre11* probe (Delmas, Shunburne *et al.* 2009), pop-outs that did not hybridise at all with the probe were selected as potential deletion mutants, shown in Figure 7.3.



**Figure 7.3 Colony lift and hybridisation of H1190 and H1191 pop-outs**

Pop-outs (*ura*-) selected for on 5FOA were patched out on Hv-Ca +thy +trp, to maintain selection for *trpA* marked deletion construct in case of merodiploidy. Four pop-out candidates that failed to hybridise with the *mre11* probe are indicated by red boxes.

The genomic sequence of *H. volcanii* was used to select appropriate restriction enzymes to digest the genomic DNA of the potential deletion mutants, which was then used in Southern hybridisation to confirm genomic deletion of *mre11-rad50* and *mre11*. A  $^{32}\text{P}$  labelled probe of *PvuI* fragment from pTA171 was used in the hybridisation (Delmas, Shunburne *et al.* 2009), shown in Figure 7.4.



**Figure 7.4 Southern blot analysis of H1190 pop-outs (H1214) and H1191 pop-outs (H1215)**

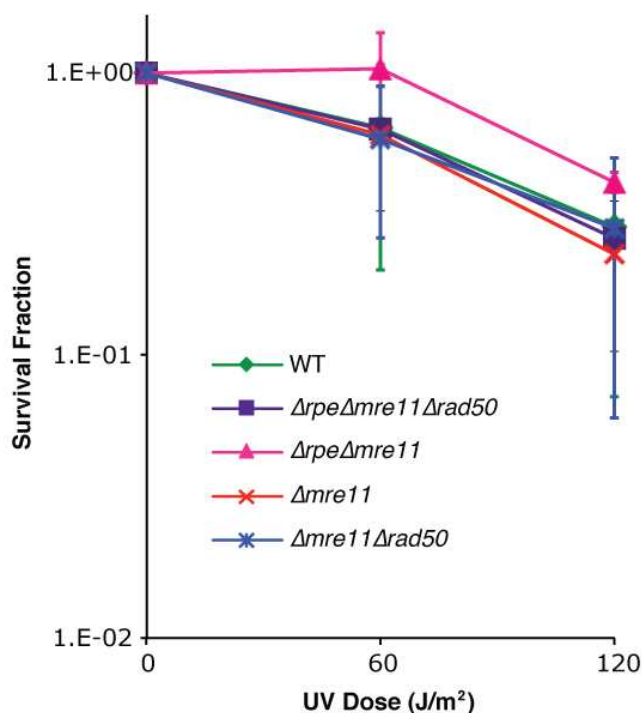
Lane M shows size marker lane. Genomic DNA was digested with *AgeI* and probed with the flanking regions of the deletion construct pTA171 *PvuI*. Genomic DNA of H678 was used a wild-type control.

### **7.1.2 DNA damage assays of $\Delta rpe \Delta mre11 \Delta rad50$ and $\Delta rpe \Delta mre11$ deletion mutants**

DNA damage assays were performed using H1214 and H1215 to examine if RPE shares a role with Mre11 and Rad50 in the processing of damaged DNA ends.

#### **UV assay**

To examine if RPE, Mre11 and Rad50 function in the nucleotide excision repair and homologous recombination pathways, the  $\Delta rpe \Delta mre11 \Delta rad50$  and  $\Delta rpe \Delta mre11$  deletion mutants were analysed in parallel with H195, and H276 and H280 that had been made previously (Delmas, Shunburne *et al.* 2009), in a UV assay. The data in Figure 7.5 shows that none of the deletion mutants are significantly sensitive to UV, indicating that RPE does not have an overlapping function with Mre11 or Rad50 in the direct processing of damaged DNA ends.

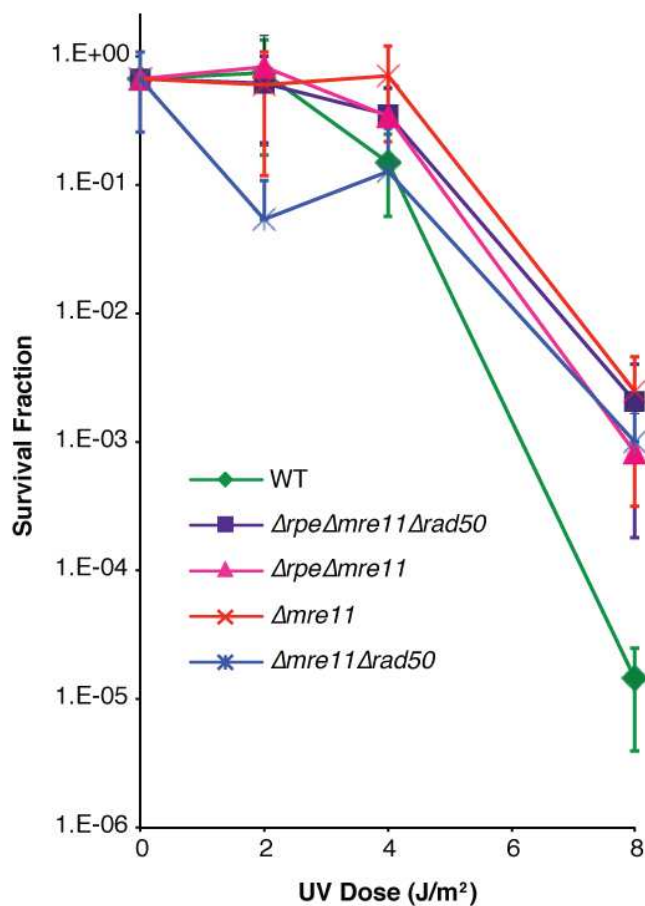


**Figure 7.5 UV sensitivity of *rpa1* deletion mutants**

The survival fraction of wild-type (WT) H195, H1214 ( $\Delta rpe \Delta mre11 \Delta rad50$ ), H1215 ( $\Delta rpe \Delta mre11$ ), H276 ( $\Delta mre11$ ) and H280 ( $\Delta mre11 \Delta rad50$ ) are plotted. The data shown is mean and standard error of three repeats.

### Hydrogen peroxide assay

To examine if RPE, Mre11 and Rad50 function in the removal of oxidative damage, the  $\Delta rpe \Delta mre11 \Delta rad50$  and  $\Delta rpe \Delta mre11$  deletion mutants were analysed in parallel with H195, H276 ( $\Delta mre11$ ) and H280 ( $\Delta mre11 \Delta rad50$ ) in a hydrogen peroxide ( $H_2O_2$ ) assay, shown in Figure 7.6. Liquid culture aliquots were incubated for an hour with 10  $\mu$ l of diluted  $H_2O_2$ , giving final concentrations of 2, 4 and 8 mM, before being plated out. The data shows none of the deletion mutants are significantly sensitive to  $H_2O_2$ , indicating that RPE does not have an overlapping function with Mre11 or Rad50 in the direct processing of oxidative-damaged DNA ends.



**Figure 7.6  $H_2O_2$  sensitivity of H1214 and H1215 deletion mutants**

The survival fraction of wild-type (WT) H195, H1214 ( $\Delta rpe \Delta mre11 \Delta rad50$ ), H1215 ( $\Delta rpe \Delta mre11$ ), H276 ( $\Delta mre11$ ) and H280 ( $\Delta mre11 \Delta rad50$ ) are plotted. The data shown is mean and standard error of three repeats.

## 7.2. $\Delta rpa1 \Delta rpa3ap$ deletion mutant

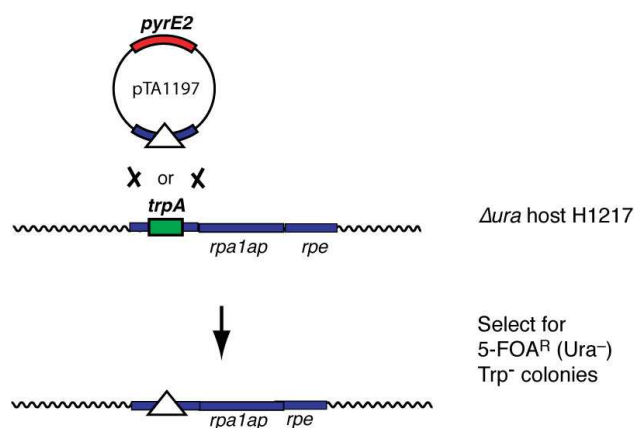
To examine if the products of the *rpa1* and the *rpa3* operons have redundant roles in DNA repair, double deletion mutants were attempted. This included the attempted generation of  $\Delta rpa1 \Delta rpa3ap$  and  $\Delta rpa1ap \Delta rpa3$  mutants, to test if one RPA can complement the absence of the other RPA, and if one associated protein (AP) can also complement the absence of the other AP. In essence, this experiment was used to examine if the function of the putative RPA:AP complex depends on specific RPA:AP interactions, or if the subunits are interchangeable.

### 7.2.1. Construction of non-*trpA* marked $\Delta rpa1$ deletion mutant

In order to use *trpA* selection when attempting to delete *rpa3ap* in an  $\Delta rpa1::trpA+$  background, the *trpA*-marked *rpa1* deletion must be replaced with a "clean" or unmarked (non-*trpA* marked) deletion.

#### Pop-in/pop-out of unmarked deletion construct

This was achieved by transforming the unmarked *rpa1* deletion construct pTA1141 (see Figure 4.8 for details) into *E. coli* dam<sup>-</sup>, generating pTA1197, which was then transformed into *H. volcanii* strain H1217 ( $\Delta rpa1$ ). Using the counter-selective pop-in/pop-out method described previously, as shown in Figure 7.7 pop-outs were generated and checked for a loss of tryptophan prototrophy by patching out on Hv-Ca +thy +ura +trp and Hv-Ca +thy +ura, to generate strain H1281.

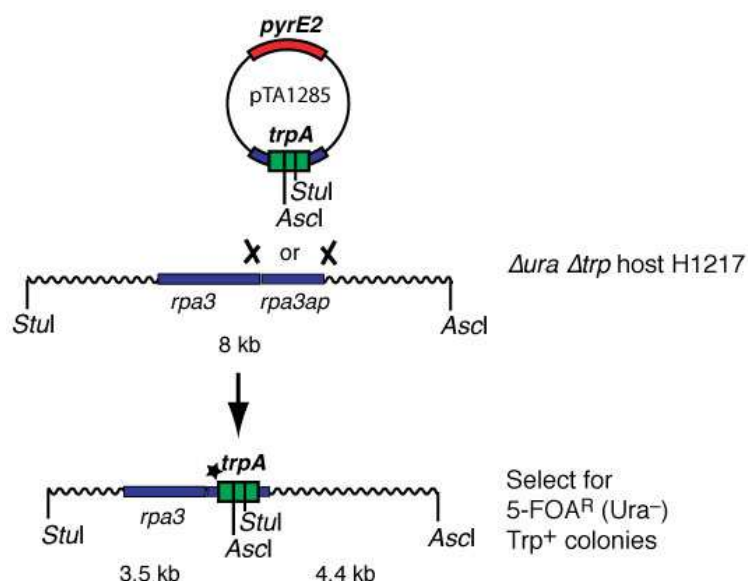


**Figure 7.7 Schematic diagram of unmarked *rpa1* deletion construction**

The deletion plasmid pTA1197 was transformed into strain H1217. Integration into the genome by homologous recombination, generated pop-in strain H1287, followed by loss of the plasmid by homologous recombination to generate pop-out strain H1281. This was followed by confirmation of *trpA*<sup>-</sup> prototrophy of the unmarked  $\Delta rpa1$  strain H1281.

### 7.2.2. Attempt to generate double $\Delta rpa1\Delta rpa3ap$ deletion mutant

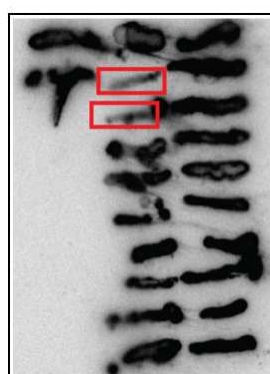
The *trpA*-marked *rpa3ap* deletion construct pTA1285 (See Chapter 5 for details) was then transformed into H1281, generating the pop-in H1473 in an attempt to delete *rpa3ap* in a  $\Delta rpa1$  background, using the counter-selective pop-in/pop-out method, shown in Figure 7.8.



**Figure 7.8. Schematic diagram of *rpa3ap* deletion construction**

The deletion plasmid pTA1285 was constructed as described in Figure 5.11. Restriction digest with *AscI* and *StuI* and the resulting fragments expected that were used in the Southern blot analysis are shown. Integration into the genome by homologous recombination generated pop-in H1473, which was followed by loss of the plasmid by homologous recombination in an attempt to generate the *rpa3ap* deletion.

Potential deletions were screened by colony hybridisation, using a  $^{32}\text{P}$  labelled *rpa3ap* probe, pop-outs that did not hybridise strongly with the probe were selected as potential deletion mutants, as shown in Figure 7.9.

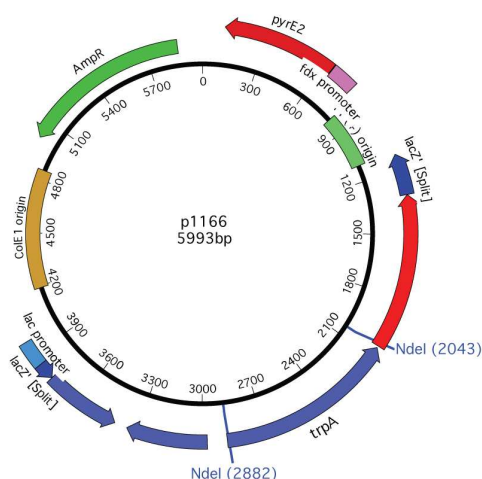


**Figure 7.9 Colony lift and hybridisation of H1473 pop-outs**

Pop-outs (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv- YPC. Two pop-out candidates that failed to hybridise strongly with the *rpa3ap* probe are indicated by red boxes

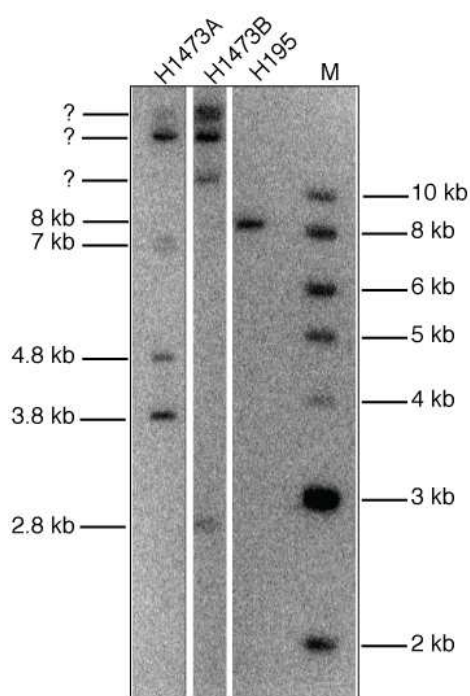


The genome sequence of *H. volcanii* was used to determine an appropriate restriction digest of the potential double deletion mutants, this was then used in Southern hybridisation, a  $^{32}\text{P}$  labelled probe of the *trpA* marker was used in the hybridisation, as shown in Figure 7.10. However, all candidates still possessed the wild-type *rpa3ap* gene.



**Figure 7.10 *trpA* Probe**

Figure shows *NdeI* restriction sites used to digest the *trpA* marker from pTA1166, which was labelled with  $^{32}\text{P}$  in Southern hybridisations.



**Figure 7.11. Southern blot analysis of H1473 pop-outs**

Lane M shows size marker. Genomic DNA was digested with *AscI/StuI* and probed with the flanking regions and the *trpA* marker shown in Figure 7.10. Genomic DNA of H195 was used as a wild-type control. An *rpa3ap* deletion should generate a 3.4 kb and 4.4 kb bands, however there are a number extra bands that cannot be accounted for, for digest details see Figure 7.8.



### 7.3. $\Delta rpa1ap$ $\Delta rpa3$ deletion mutant

To examine if the products of the *rpa1* and the *rpa3* operons have redundant roles in DNA repair, double deletions were attempted. This included the attempted generation of  $\Delta rpa1ap$   $\Delta rpa3$  mutants, to test if one RPA1 can complement the absence for RPA3, and if RPA1AP can also complement for RPA3AP.

#### 7.3.1. Construction of *rpa1ap* deletion mutant

In order replace the *trpA*-marked *rpa1ap* deletion with an unmarked *rpa1ap* deletion, an unmarked *rpa1ap* deletion construct was made, as shown in Figure 7.12.

#### Design of primers

The upstream region of *rpa1ap* was amplified from the genomic *rpa1* clone pTA937 by touchdown PCR using the primers shown in Table 7.1, introducing external *KpnI* and internal *SphI* restriction sites.

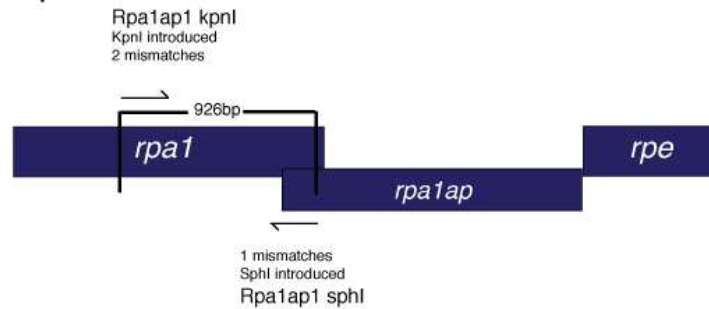
Primer	Sequence (5'-3')	Site inserted
Rpa1ap1 kpnI	CCGCGAGTGGtACcGCAAGCCCG	<i>KpnI</i>
Rpa1ap1 sphI	GCGATTTCCCGCaTGCCGACGACCG	<i>SphI</i>

**Table 7.1 Design of upstream deletion primers for unmarked *rpa1ap* deletion**

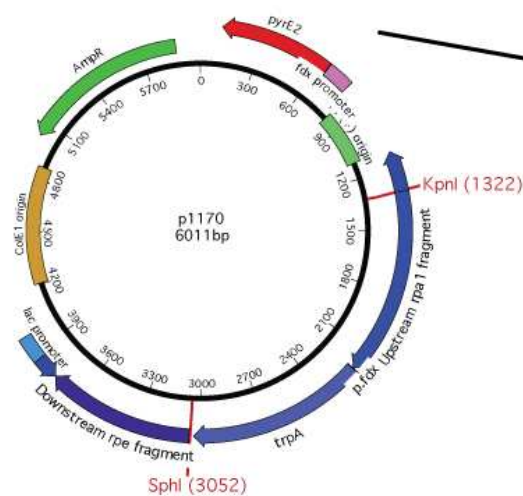
The sequence for each deletion primer is shown with the base changes, (in lower case), made to introduce a novel restriction site (underlined), to allow cloning of the upstream region of *rpa1ap*. The Rpa1ap1 kpnI primer was used in the previous construction of the *rpa1 trpA*-marked deletion (See Chapter 4.2.2).

## Construction of unmarked *rpa1ap* deletion plasmid

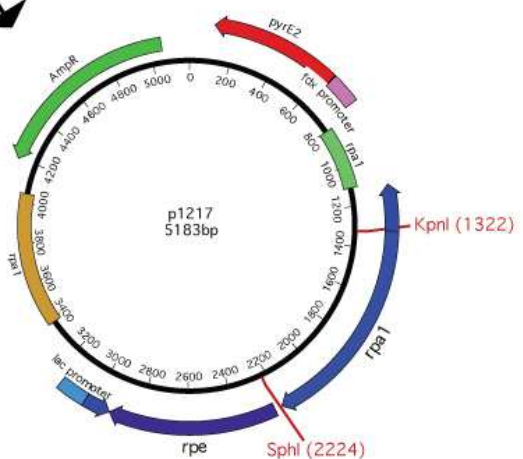
### A) RPA1 operon



### B) Plasmid pTA1170 containing *trpA*-marked *rpa1ap* deletion construct



### C) Plasmid pTA1217 showing unmarked *rpa1ap* deletion



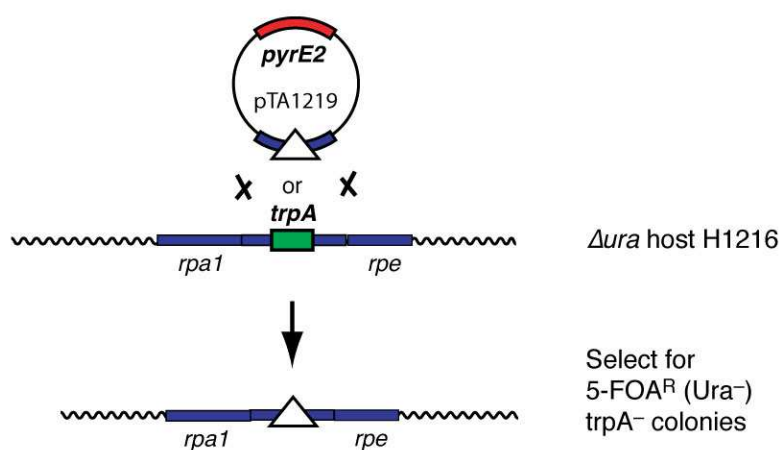
- D)
1. Amplification of upstream region using primers shown
  2. *KpnI* and *SphI* digest of PCR product and pTA1170
  3. Insert upstream region at *KpnI* and *SphI* sites in pTA1170
  4. *E. coli* dam<sup>+</sup> transformants selected for using ampicillin resistance

**Figure 7.12 Construction of unmarked *rpa1ap* deletion plasmid**

Figure shows the cloning technique used to generate the unmarked deletion construct of *rpa1ap*, the binding sites of the primers are shown A) the restriction sites in the vector pTA1170 used to insert the upstream ligation B) and the resulting deletion construct pTA1219 C). A basic sequence of construction is shown in D).

### Pop-in/pop-out of unmarked deletion construct

The unmarked *rpa1ap* deletion construct pTA1217 was transformed into *E. coli dam*-generating pTA1219, which was then transformed into *H. volcanii* H1216 generating the pop-in H1286, as shown in Figure 7.13. This pop-in strain was then used to generate the pop-out and unmarked *rpa1ap* deletion strain H1280.

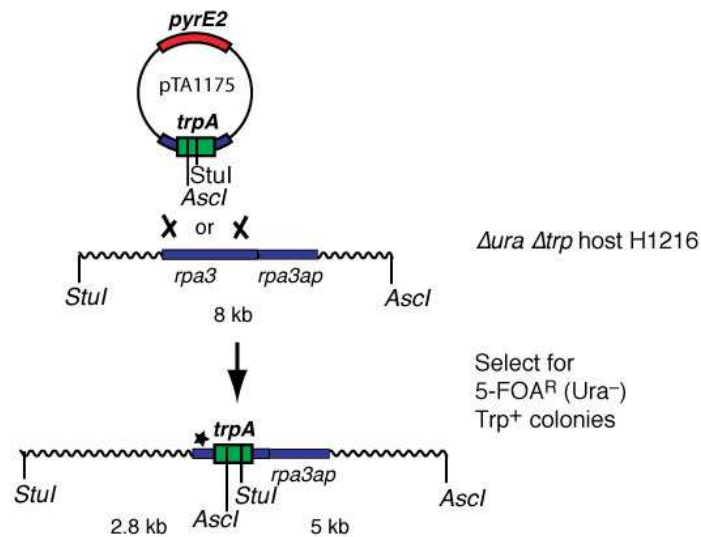


**Figure 7.13. Schematic diagram of non-*trpA* marked *rpa1ap* deletion construction**

The deletion plasmid pTA1219 was constructed as described in Figure 7.12. Integration into the genome by homologous recombination generated pop-in strain H1286, followed by loss of the plasmid by homologous recombination to generate pop-out strain H1280, which was followed by confirmation of *trpA*<sup>-</sup> prototrophy of the unmarked *Δrpa1ap* strain H1280.

### 7.3.2. Attempt to generate double *Δrpa1apΔrpa3* deletion mutant

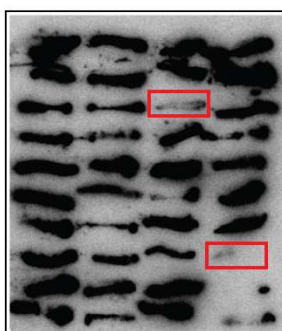
The *trpA*-marked *rpa3* deletion construct pTA1175 (See Chapter 5 for details) was then transformed into the unmarked *Δrpa1ap* strain H1280, generating the pop-in H1326 in an attempt to delete *rpa3* in an *Δrpa1ap* background, using the counter-selective pop-in/pop-out method shown in Figure 7.14.



**Figure 7.14 Schematic diagram of *rpa3* deletion construction**

The deletion plasmid pTA1175 was constructed as described in Figure 5.6. Restriction digest with *AscI* and *StuI* and the resulting fragments expected that were used in the Southern blot analysis are shown. Integration into the genome by homologous recombination generated pop-in strain H1326, followed by loss of the plasmid by homologous recombination in an attempt to generate the *rpa3* deletion.

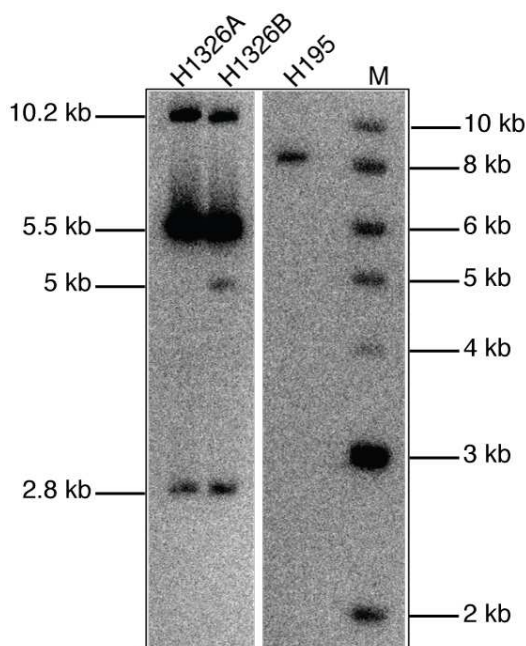
Potential deletions were screened by colony hybridisation, using a  $^{32}\text{P}$  labelled *rpa3* probe (see Figure 5.1), pop-outs that did not hybridise strongly with the probe were selected as potential deletion mutants.



**Figure 7.15 Colony lift and hybridisation of H1326 pop-outs**

Pop-outs (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv- YPC. Two pop-out candidates that failed to hybridise strongly with the *rpa3* probe are indicated by red boxes

The genome sequence of *H. volcanii* was used to determine an appropriate restriction digest of the potential double deletion mutants, this was then used in Southern hybridisation, a  $^{32}\text{P}$  labelled probe of the *trpA* marker was used in the hybridisation, as shown in Figure 7.16. However all mutants still possessed the wild-type *rpa3* gene.



**Figure 7.16 Southern blot analysis of H1326 pop-outs**

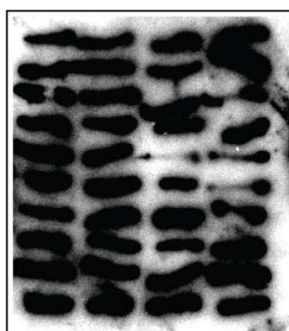
Lane M shows the size marker. Genomic DNA was digested with *AscI* and *StuI* and probed with the *trpA* marker shown in Figure 7.10. Genomic DNA of H195 was used as a wild-type control. Expected band sizes should have been 2.8 kb and 5 kb, these are present in H1326 but the predominant band is 5.5 kb that cannot be accounted for, suggesting *rpa3* has not been deleted. The presence of a band of 10.2 kb cannot be accounted for. See Figure 7.14 for further digest details.

## 7.4. $\Delta rpa1ap$ $\Delta rpa3ap$ deletion mutant

### 7.4.1. Attempt to generate double $\Delta rpa1ap$ $\Delta rpa3p$ deletion mutant

To examine if the associated proteins of RPA1 and RPA3 are required for RPA function, an attempt to make a double  $\Delta rpa1ap$   $\Delta rpa3ap$  was made. The *rpa3ap* deletion construct pTA1185 was transformed into H1280 (unmarked *rpa1ap* deletion strain) generating the pop-in strain H1492. The counter-selective method was used to generate pop-outs as shown in Figure 4.6.

Potential deletions were screened by colony hybridisation, using  $^{32}\text{P}$  labelled *rpa3ap* probe (see figure 5.1), however all pop-outs were merodiploid and hybridised with the probe, shown Figure 7.17.



**Figure 7.17 Colony lift and hybridisation of H1492 pop-outs**

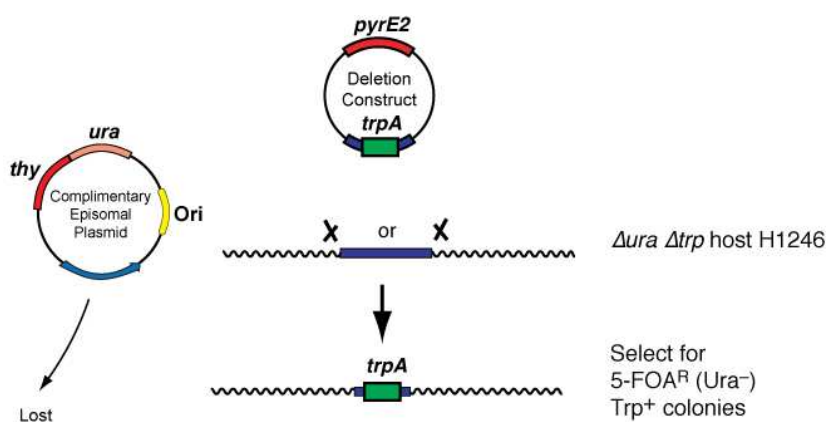
Pop-outs (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv-Ca +thy. All pop-outs hybridised with the *rpa3ap* probe and were therefore merodiploid.

## 7.5. $\Delta rpa1$ operon $\Delta rpa3$ operon deletion mutants

As shown in the phylogenetic analysis in Chapter 3, *rpa1* and *rpa3* are closely related indicating that there is an element of redundancy. Both the  $\Delta rpa1$  operon and the  $\Delta rpa3$  operon mutants did not display a severe growth defect. Therefore to establish if they are required for cell growth and survival, or if RPA2 alone is sufficient, a double  $\Delta rpa1$  operon  $\Delta rpa3$  operon mutant was made.

### 7.5.1 Construction *rpa1* operon complementation plasmid

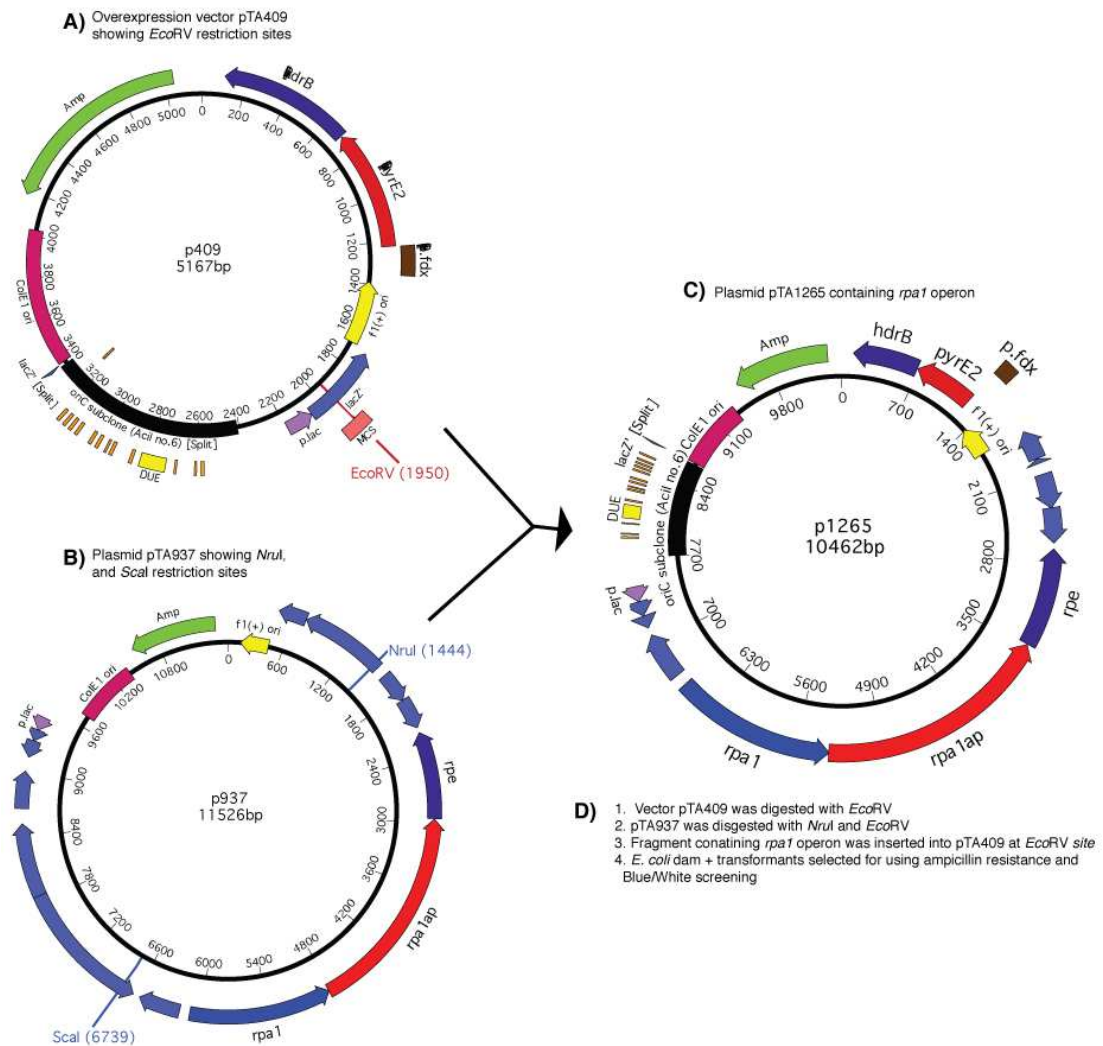
The failure to make the  $\Delta rpa1 \Delta rpa3$  and the  $\Delta rpa1 \Delta rpa3$  mutants indicates that generating a double  $\Delta rpa1$  operon  $\Delta rpa3$  operon mutant would be difficult or impossible. To overcome this an episomal plasmid providing RPA1 and RPA1AP *in trans* was constructed, as shown in Figure 7.19, and used to generate a  $\Delta rpa1$  operon  $\Delta rpa3$  operon mutant. This technique is shown in Figure 7.18, and has been used previously to successfully generate  $\Delta radA$  mutants in a  $\Delta mre11 \Delta rad50$  background (Delmas, Shunburne *et al.* 2009).



**Figure 7.18 Pop-in/pop-out method using complimentary episomal plasmid**

Counter-selective pop-in/pop-out method shown, using a complementary episomal plasmid to provide *in trans* expression of the gene to be deleted. Adapted from (Leigh, Albers *et al.* 2011).

An episomal plasmid containing the *rpa1* operon for complementation during pop-out process was transformed into strain H1246 (unmarked *rpa1* operon deletion), to generate the strain H1282.



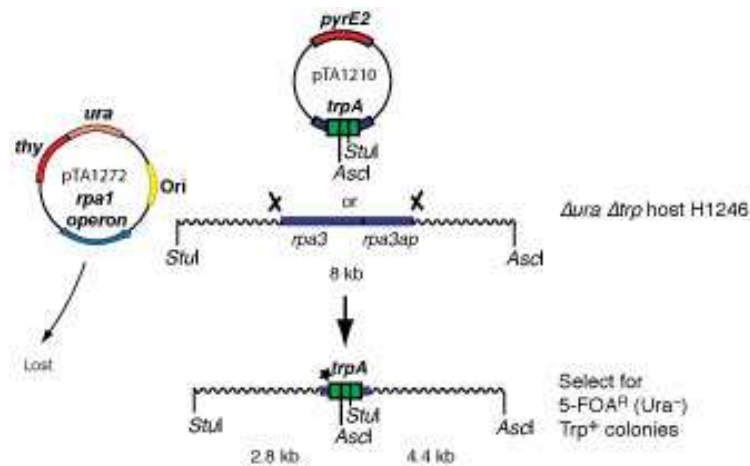
**Figure 7.19 Construction of the *rpa1* operon complementation plasmid**

A) Shuttle vector pTA409 showing *EcoRV* site B) *rpa1* operon clone pTA937 showing *NruI* and *ScaI* restriction sites C) Complementation plasmid pTA1265 containing *rpa1* operon, generated by ligation of *NruI* and *ScaI* sites to a *EcoRV* site.

### 7.5.2 Generation of double $\Delta rpa1$ operon $\Delta rpa3$ operon deletion mutant

The *trpA*-marked *rpa3* operon deletion construct pTA1210 was transformed into strain H1282 generating the pop-in strain H1390. This was then used to generate the pop-out strain H1472 using the counter-selective method as shown in Figure 7.20.

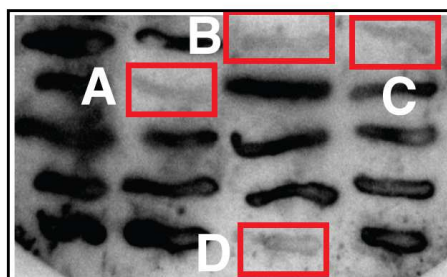




**Figure 7.20 Schematic diagram of *rpa3* deletion construction**

The *trpA*-marked *rpa3* operon deletion plasmid pTA1210 and the *rpa1* operon complementation plasmid were constructed as shown in Figures 5.15 and 7.19, respectively. Restriction digest with *AscI* and *StuI* and the resulting fragments expected that were used in the Southern blot analysis are shown. Integration into the genome by homologous recombination generated pop-in strain H1390, followed by loss of the integrated plasmid by homologous recombination to generate the double *rpa1* operon *rpa3* operon deletion strain H1472.

Potential deletions were screened by colony hybridisation using a  $^{32}\text{P}$  labelled *rpa3* probe, pop-outs that did not hybridise to the probe were selected as potential deletion mutants, as shown in Figure 7.21.



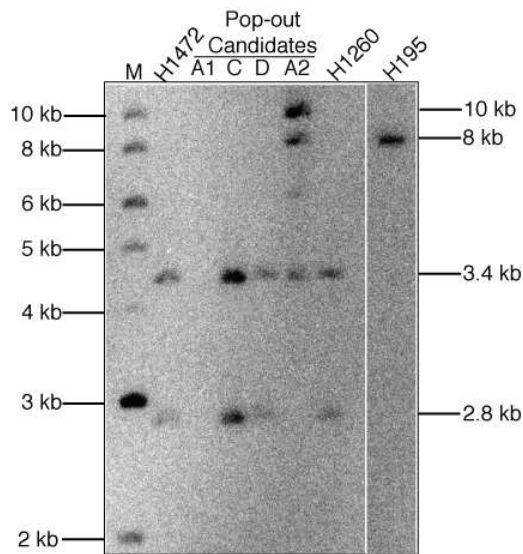
**Figure 7.21 Colony lift and hybridisation of H1390 pop-outs**

Pop-outs (*ura*<sup>-</sup>) selected for on 5FOA were patched out on Hv- YPC. Four pop-out candidates (A-D) that failed to hybridise with *rpa1* probe are indicated by red boxes.

The genomic sequence of *H. volcanii* was used to determine an appropriate restriction digest to screen the four potential deletion mutants, this was then used in a Southern hybridisation to confirm genomic deletion of the *rpa1* operon, shown in Figure 7.22. Pop-out candidate A grew poorly in selective Hv-Ca +thy +ura media (A1) so a genomic preparation was prepared from candidate A grown in Hv-YPC +thy (A2); this candidate turned out to be merodiploid as shown in Figure 7.21. A  $^{32}\text{P}$  labelled



probe of the *trpA* was used in the hybridisation, shown in Figure 7.22, the pop-out selected (B) is labelled as H1472.

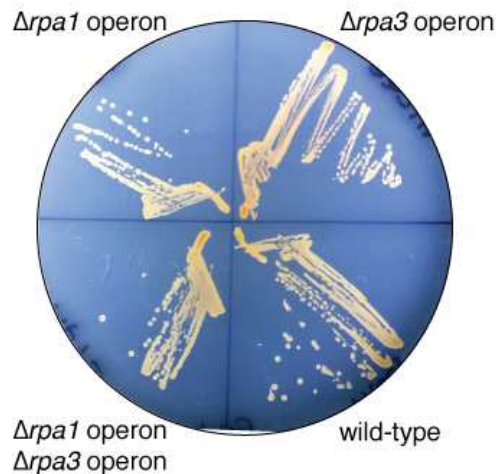


**Figure 7.22 Southern blot analysis of H1245 pop-outs (H1472)**

Lane M shows size marker lane. Genomic DNA was digested with *AscI* and *StuI* and probed with the *trpA* marker shown in figure 7.10. Genomic DNA of H195 was used as a wild-type control. *rpa3* operon deletion generates 2.8 kb and 3.4 kb bands, while wild-type *rpa3* operon gives a 8 kb band. The presence of the 10 kb band in the (merodiploid) candidate A2 cannot be accounted for. See Figure 7.20 for further restriction digest details.

### **Growth comparison of the double operon deletion mutant**

The double *rpa1 rpa3* operon strain showed a slight growth defect compared to the wild-type, single *rpa1* operon deletion and the single *rpa3* operon deletion strains, as shown in Figure 7.23.



**Figure 7.23 Double *rpa1 rpa3* operon growth analysis.**

Figure shows the growth of the double *rpa1 rpa3* operon deletion (H1472) compared to that of wild-type (H195), *rpa1* operon deletion (H1246) and *rpa3* operon deletion (H1260).

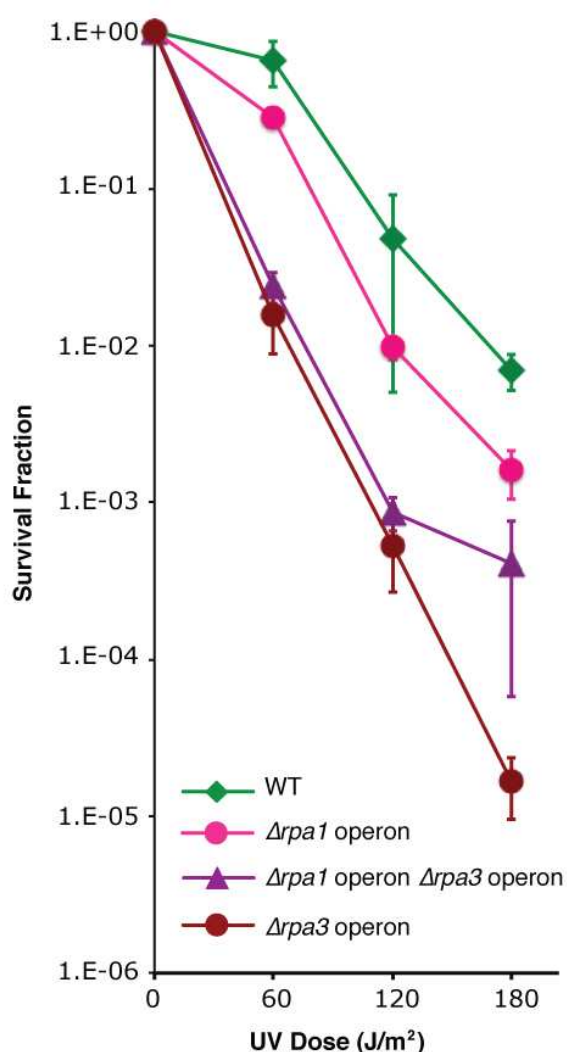
### **7.5.2. DNA damage assays of deletion mutants**

Previous work in Chapters 4 and 5 had established that the products of the *rpa1* operon do not play a significant role in DNA repair, while the products of the *rpa3* operon do. To examine if the products of the *rpa1* and the *rpa3* operon have

overlapping roles in DNA repair or if they act in different DNA repair pathways, DNA damage assays were performed with the double operon mutant.

### UV DNA damage assays

UV assays were performed to examine if the RPA1 and RPA3 operon have overlapping roles in the repair of UV-induced DNA damage, or if they have independent roles acting in different pathways or at different stages in the repair of UV-induced DNA damage. The *rpa1* operon *rpa3* operon double deletion mutant was analysed in parallel with H195 in a UV assay, shown in Figure 7.24. The data shows that the double *rpa1 rpa3* operon mutant is significantly sensitive to UV, and that the single  $\Delta rpa3$  operon mutant and the double  $\Delta rpa1$  operon  $\Delta rpa3$  operon mutant show the same degree of UV sensitivity (excluding the 180 J/m<sup>2</sup> datapoint).

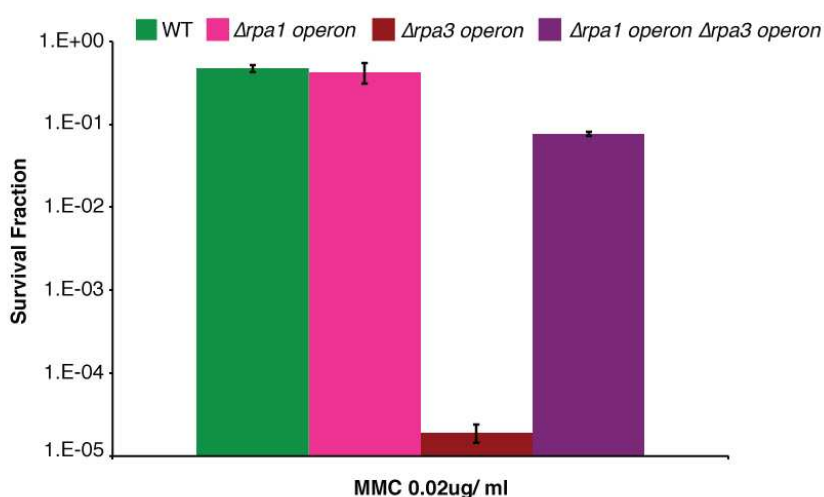


**Figure 7.24 UV sensitivity of double *rpa1 rpa3* operon deletion mutants**

The survival fraction of wild-type (WT) H195, H1246 ( $\Delta rpa1$  operon), H1472 ( $\Delta rpa1$  operon *rpa3* operon) and H1260 ( $\Delta rpa3$  operon) is plotted. The data shown is the mean and standard error of three repeats.

### MMC DNA damage assays

MMC assays were performed to examine if the RPA1 and RPA3 operon have overlapping roles in the repair of MMC-induced DNA damage, or if they have independent roles acting in different pathways or at different stages in the repair of MMC-induced DNA damage. The *rpa1* operon and *rpa3* operon deletion mutant were analysed in parallel with H195 in a MMC assay, shown in Figure 7.25, Hv-YPC +thy plates containing 0.02 µg/ml MMC were used. The data shows that the double  $\Delta rpa1$  operon  $\Delta rpa3$  operon deletion mutant is significantly sensitive to MMC, but not as sensitive as the single  $\Delta rpa3$  operon mutant.



**Figure 7.25 MMC sensitivity of *rpa1* deletion mutants**

The survival fraction of the wild-type (WT) H195, H1246 ( $\Delta rpa1$  operon), H1260 ( $\Delta rpa3$  operon) and H1472 ( $\Delta rpa1$  operon  $\Delta rpa3$  operon) are plotted. The data shown is the mean and standard error of three repeats.

## 7.6 Discussion

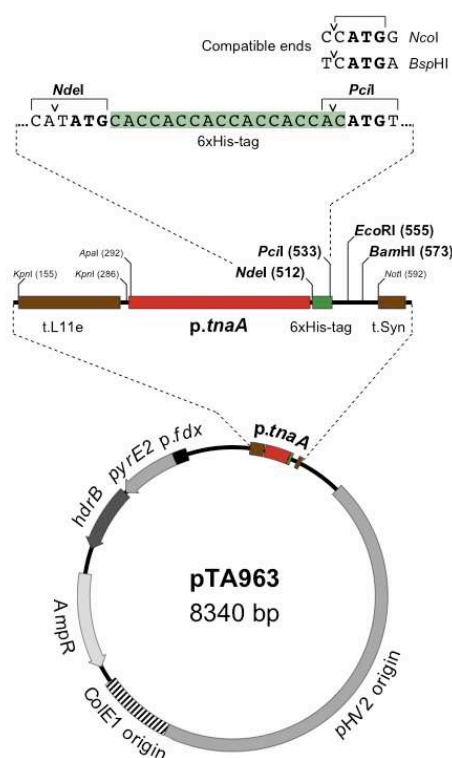
To examine if RPE has a redundant role with Mre11 in the processing of damaged DNA ends, the double  $\Delta rpe \Delta mre11$  and the triple  $\Delta rpe \Delta mre11 \Delta rad50$  deletion mutants were made. The ease at which the triple  $rpe$ ,  $mre11$  and  $rad50$  and the double  $rpe$  and  $mre11$  deletion mutants were made suggests they do not have an overlapping role in the processing of damaged DNA ends. This is supported by the UV and hydrogen peroxide DNA damage assays.

In order to make the double  $rpa1 rpa3$  operon deletion, it was necessary to use an episomal plasmid to provide *in trans* expression of the  $rpa1$  operon. Such a method of complementation was also necessary to make certain  $\Delta radA$  mutants (Delmas, Shunburne *et al.* 2009). However unlike the  $\Delta radA$  mutants, which displayed a severe growth defect and significant UV sensitivity, only a slight growth defect is seen in the double  $rpa1 rpa3$  operon deletion mutant. This demonstrates that the products of the two RPA operons aid cell growth and duplication but are not essential. In addition, it is important to consider the possibility that the milder sensitivity to UV and MMC, compared to that of the single  $rpa3$  operon deletion mutants, could indicate that the products of the  $rpa1$  and  $rpa3$  operon act in the same DNA repair pathway.

In combination with a decrease in sensitivity to DNA damage compared to that of the single  $\Delta rpa3$  operon, this raises the question of whether the double  $\Delta rpa1 \Delta rpa3$  operon mutant possesses a suppressor mutation. A suppressor mutation would explain why the double operon mutant was difficult to make, requiring *in trans* expression of the  $rpa1$  operon. However, once the double deletion mutant had been generated and the suppressor mutation had arisen, there was a lack of a growth defect or severe DNA damage sensitivity. Perhaps the suppressor mutation lies within  $rpa2$ , allowing it to compensate for the absence of RPA1 and RPA3. An increased resistance to ionizing irradiation is seen in the *Halobacterium* sp. NRC-1, due to the upregulation of expression of two RPAs (DeVeaux, Muller *et al.* 2007). This suggests that mild sensitivity to DNA damage and absence of a growth defect of the *H. volcanii* double  $rpa$  operon mutant could be the result of upregulation of  $rpa2$ , which could be identified by sequencing  $rpa2$  from the double deletion mutant.

## Chapter 8: Histidine-tagged RPA and associated proteins

To examine if RPA1 and RPA3 interact with their respective associated proteins (APs), and if RPA2 interacts with RPA1 and/or RPA3, the genes encoding the three RPAs and two APs were inserted into an overexpression plasmid pTA963 containing a hexahistidine-tag, shown in Figure 8.1 (Allers, Barak *et al.* 2010). All of plasmid constructs, except those containing *rpa2*, were found to be toxic in *E. coli* XL-1. Consequently they were constructed in *H. volcanii* H1209 ( $\Delta mrr$ ), which is deleted for the Mrr restriction enzyme that cuts at methylated GATC sequences and therefore does not require dam– DNA for efficient transformation (Allers, Barak *et al.* 2010).

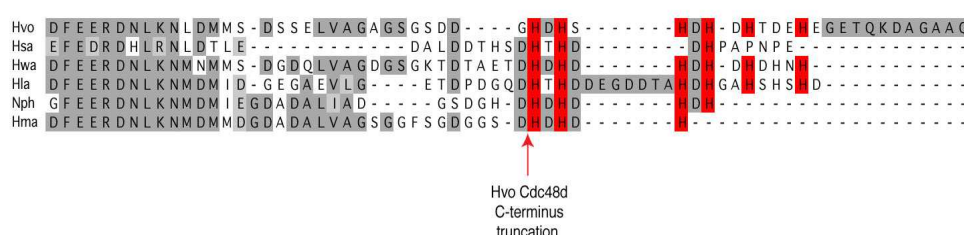


**Figure 8.1 Overexpression vector pTA963**

The overexpression vector pTA963 (Allers, Barak *et al.* 2010) containing a hexahistidine-tag (6-His-tag) under the tryptophanase promoter (*p.tnaA*), allowing inducible expression of a histidine-tagged gene. pTA963 features the selectable markers *pyrE2*, *hdrB* and ampicillin resistance.

The histidine-tagged constructs were then transformed into *H. volcanii* H1424, which features two key modifications that reduce the co-purification of naturally histidine-rich proteins with hexahistidine-tagged recombinant proteins. The *pitA* gene of *H. volcanii*, which contains a histidine-rich linker region, was previously replaced with an orthologue from *Natronomonas pharaonis* that is not histidine-rich (Allers, Barak *et al.* 2010). The *cdc48d* gene, encoding a protein that features a histidine-rich C-terminus, was replaced with orthologues from *Halobacterium salinarum* and

*Halobacterium marismortui*. The latter proteins have fewer histidines in their C-termini (Figure 8.2), however both gene replacements resulted in cellular growth defects. Instead, the C-terminus of the *H. volcanii* Cdc48d protein was deleted (shown in Figure 8.2). To accomplish this, the truncated *H. volcanii cdc48d* gene was used to replace the previously-generated *H. salinarum cdc48d* gene replacement by standard pop-in/pop-out techniques, this was then confirmed by (absence of) colony hybridisation using a probe of the *H. salinarum cdc48d* gene (T. Allers unpublished work).

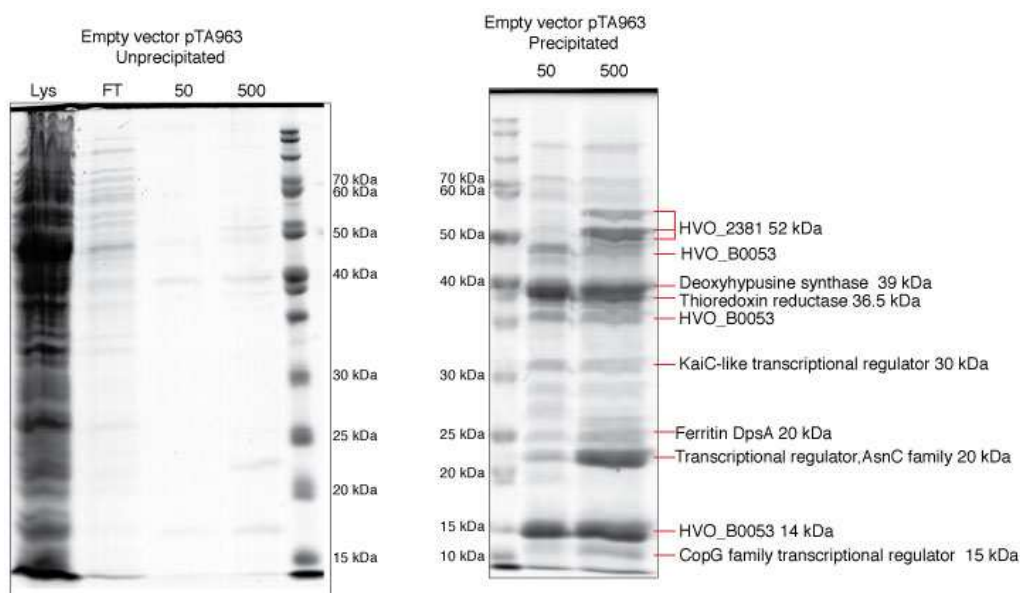


**Figure 8.2 The Cdc48d C-terminus**

Alignment of the Cdc48d C-terminus from haloarchaea, including *H. volcanii* (Hvo), *H. salinarum* (Hsa), *N. pharaonis* (Nph) and *H. marismortui* (Hma).

Tryptophan-induced overexpression of the histidine-tagged proteins was followed by cell disruption by sonication and overnight incubation of the cell extract with nickel-charged Sepharose beads, which the histidine-tagged proteins and any strongly interacting proteins will bind to. This mixture was then run through a gravity column, using the nickel-charged beads to bind the histidine-tagged and interacting proteins, thereby purifying them from the rest of the cellular proteins. After elution from the nickel column using an imidazole step gradient (50 and 500 mM imidazole), the protein sample was run out on a SDS-PAGE gel.

An empty vector control, strain H1430 containing pTA963, was used with each set of pull-downs and is shown in Figure 8.3. The bands visible are proteins that naturally feature histidine and aspartic acid residues, but not to the extent of PitA and Cdc48d.



**Figure 8.3 Empty vector pull-down**

Empty vector mock pull-down using pTA963. 50 and 500 mM imidazole was used for elution from the nickel column. Unprecipitated elutions are shown on the left, to enhance visualisation the protein elution were precipitated, as shown on the right. Annotated bands were identified by mass spectrometry (data not shown).

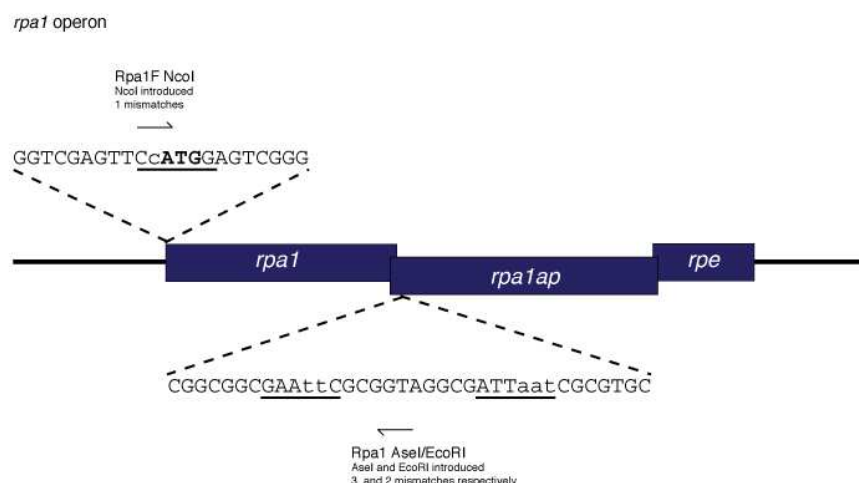
## 8.1. Construction of histidine-tagged *rpa1* with *rpa1ap*

To examine if RPA1 and RPA1AP interact with each other, an overexpression plasmid containing histidine-tagged *rpa1* and *rpa1ap* (in an operon) was constructed.

### 8.1.1. Construction of histidine-tagged *rpa1*

#### Design of primers

The *rpa1* gene was amplified from the genomic *rpa1* clone pTA937 by touchdown PCR using the primers shown in Table 4.2.1, introducing N-terminal *Nco*I and C-terminal *Eco*RI restriction sites. The restriction site *Nco*I was chosen, since it contains ATG followed by a G, to avoid changing the nucleotide G after the start codon (ATG) and altering the reading frame of the operon (see Figure 8.1).



**Figure 8.4 Introduction of *NcoI* and *EcoRI* sites**

The introduction of N-terminal *NcoI* and C-terminal *EcoRI* restriction sites maintains the reading frame of *rpa1*. Start codons shown in bold.

Primer	Sequence (5'-3')	Site inserted
Rpa1F <i>NcoI</i>	CCCGACTC <b>cATG</b> GAACCTCGACC	<i>NcoI</i>
Rpa1 <i>AseI</i> / <i>EcoRI</i>	CGGCGGCGAA <u>ttCGCGGTAGGCGATTaa</u> <u>ttCGCGTGC</u>	<i>EcoRI</i> / <i>AseI</i>

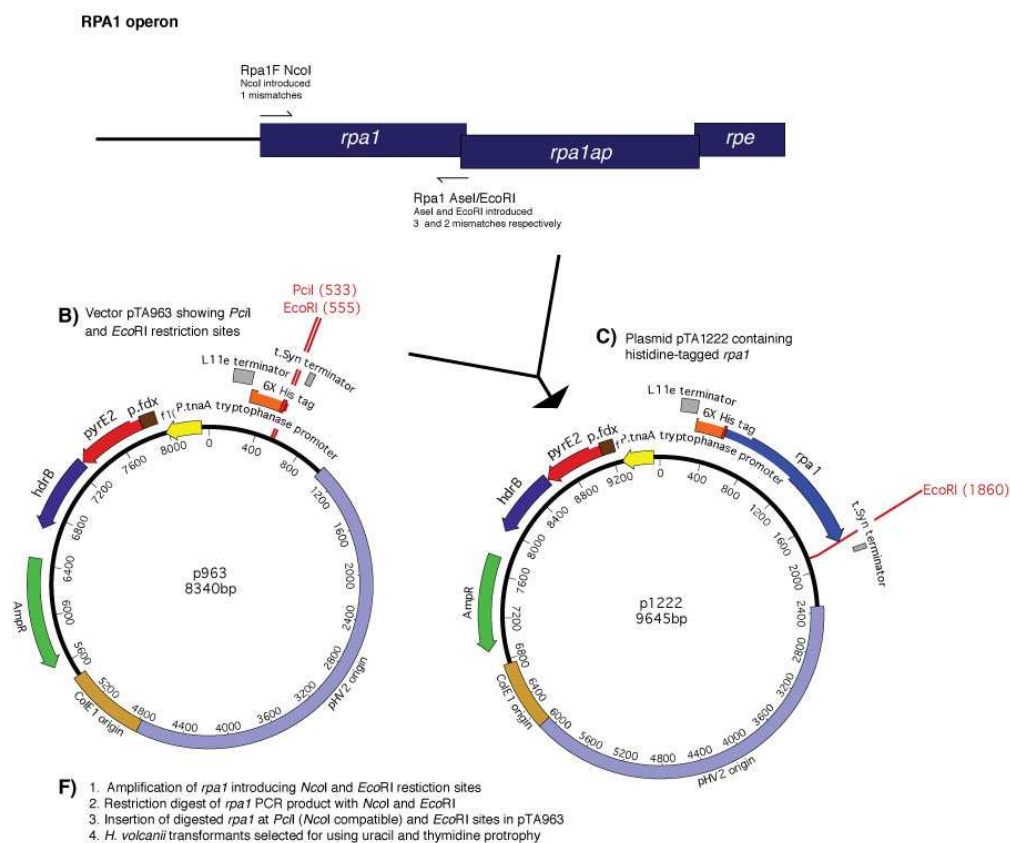
**Table 8.2 Design of primers to insert *rpa1* into pTA963**

The sequence for each primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of *rpa1* into pTA963. Start codon of the *rpa1* is shown in bold.

### Construction of histidine-tagged *rpa1* plasmid

The PCR product of *rpa1* was gel purified, digested with *NcoI* and *EcoRI* and ligated into the plasmid pTA963 using the *PciI* and *EcoRI* sites, and was transformed into *E. coli* XL-1 strain, as shown in Figure 8.5 (see also Figure 8.1). Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1222 construct.





**Figure 8.5 Construction of the histidine-tagged *rpa1* plasmid**

A) Binding sites of the primers are shown, the introduction of *AseI* site is not relevant here and will be discussed later B) the restriction sites in the overexpression vector used to insert amplified *rpa1* C) the resulting overexpression construct containing *rpa1* pTA1222. D) A basic sequence of construction.

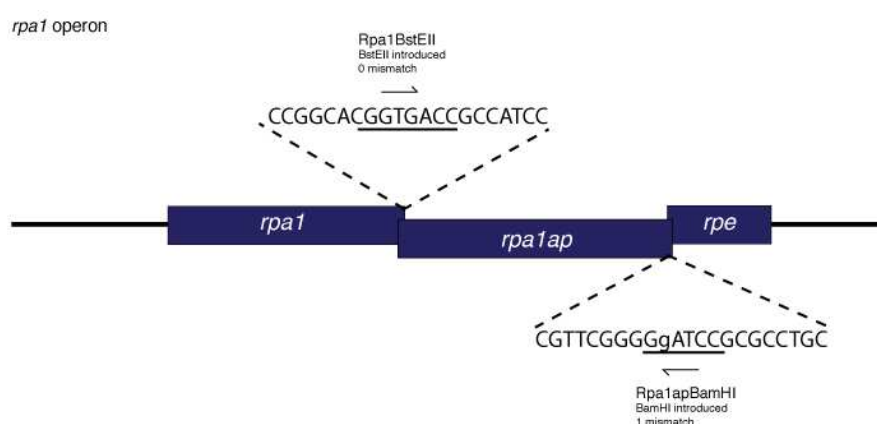
### 8.1.2. Insertion of untagged *rpa1ap* downstream of histidine-tagged *rpa1*

In order to maintain protein stoichiometry of RPA1 and RPA1AP in the cell, both *rpa1* and *rpa1ap* must be co-expressed from the same overexpression construct. Therefore *rpa1ap* was inserted downstream of the histidine-tagged *rpa1*.

### Design of primers

The *rpa1ap* gene was amplified from the genomic *rpa1* operon clone pTA937 by touchdown PCR using the primers shown in Table 8.3, featuring an N-terminal *BstEII* and C-terminal *BamHI* restriction sites. The naturally-occurring restriction site *BstEII*

located in the *rpa1* gene was chosen since it does not involve any base-pair changes and maintains the reading frame of the *rpa1* operon.



**Figure 8.6 Introduction of internal *BstEII* and *BamHI* sites**

A) The use of N-terminal *BstEII* and C-terminal *BamHI* restriction sites maintains the reading frame of *rpa1ap*. Start codons shown in bold.

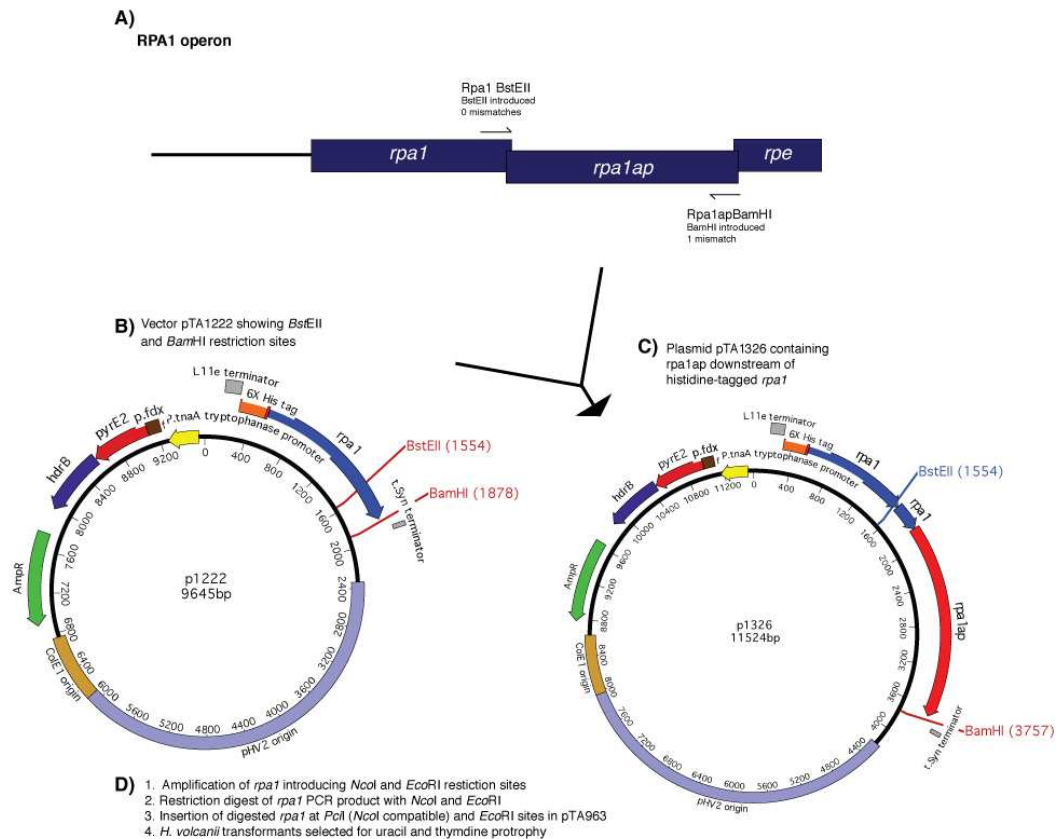
Primer	Sequence (5'-3')	Site inserted
Rpa1BstEII	CCGGCACGGTGACCGCCATCC	<i>BstEII</i>
Rpa1apBamHI	CGTTCGGGGgATCCGCGCCTGC	<i>BamHI</i>

**Table 8.3 Design of primers to insert *rpa1ap* gene into pTA1222**

The sequence for each primer is shown with the base changes, (in lower case), restriction site, (underlined), to allow cloning of *rpa1ap* into pTA1222. Start codon of the *rpa1ap* is shown in bold.

### **Insertion of *rpa1ap* downstream of histidine-tagged *rpa1***

The PCR product of *rpa1ap* was gel purified, digested with *BstEII* and *BamHI* and ligated into the plasmid pTA1222 using the *BstEII* and *BamHI* sites, and was transformed into *E. coli* XL-1 strain, as shown in Figure 8.7. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1326 construct.



**Figure 8.7 Insertion of untagged *rpa1ap* into pTA1222**

A) Binding sites of the primers are shown B) the restriction sites in the overexpression vector used to insert amplified *rpa1ap* C) the resulting overexpression construct containing histidine-tagged *rpa1* and untagged *rpa1ap* pTA1326. D) A basic sequence of construction.

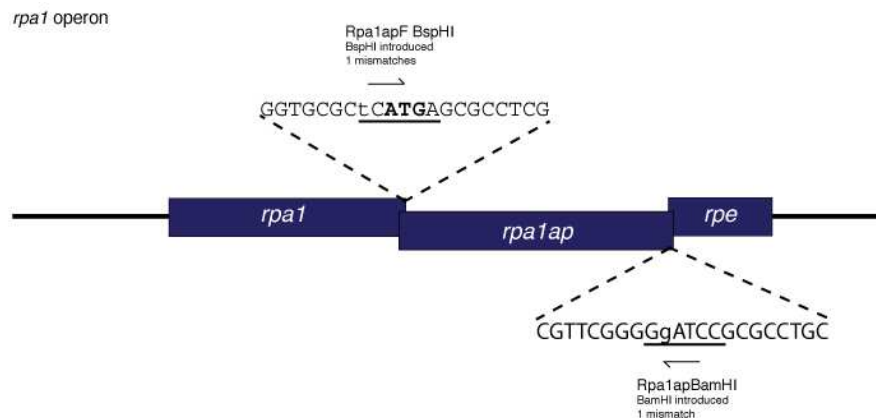
## 8.2. Construction of *rpa1* with histidine-tagged *rpa1ap*

To examine RPA1AP interacts with RPA1, an over expression plasmid containing *rpa1* and histidine-tagged *rpa1ap* (in an operon) was constructed.

### 8.2.1. Construction of histidine-tagged *rpa1ap*

#### Design of primers

The *rpa1ap* gene was amplified from the genomic *rpa1* operon clone pTA937 by touchdown PCR using the primers shown in Table 8.4, introducing a N-terminal *BspHI* and C-terminal *BamHI* restriction sites. The restriction site *BspHI* was chosen, since it contains ATG followed by an A, to avoid changing the nucleotide A after the start codon (ATG) and altering the reading frame of the operon (see Figure 8.1).



**Figure 8.8 Introduction of *Bsp*HI and *Bam*HI sites**

The introduction of N-terminal *Bsp*HI and C-terminal *Bam*HI restriction sites maintains the reading frame of *rpa1ap*. Start codons shown in bold.

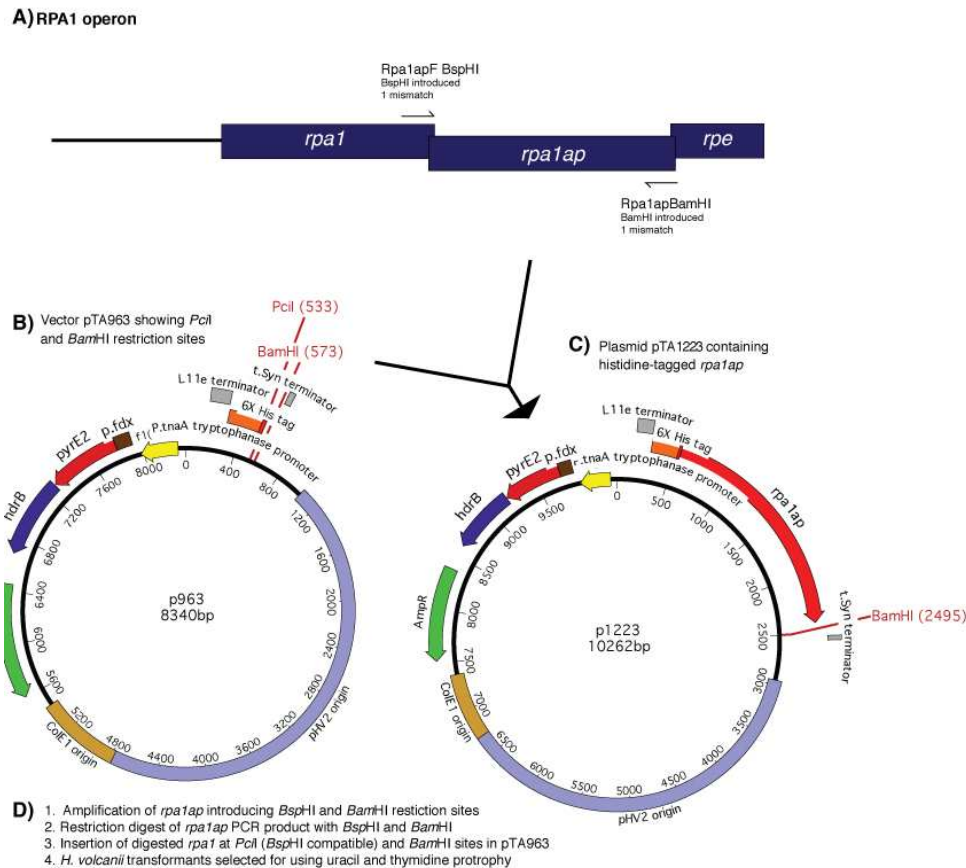
Primer	Sequence (5'-3')	Site inserted
Rpa1apF BspHI	GGTGCGCt <b>CATG</b> AGCGCCTCG	<i>Bsp</i> HI
Rpa1apBamHI	CGTTCGGGGg <b>ATCC</b> GCGCCTGC	<i>Bam</i> HI

**Table 8.4 Design of primers to insert *rpa1ap* into pTA963**

The sequence for each primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of *rpa1ap* into pTA963. Start codon of the *rpa1ap* is shown in bold.

### Construction of histidine-tagged *rpa1ap* plasmid

The PCR product of *rpa1ap* was gel purified, digested with *Bsp*HI and *Bam*HI, and ligated into the plasmid pTA963 using the *Pci*I and *Eco*RI sites, followed by transformation into *E. coli* XL-1 strain, as shown in Figure 8.9 (see also Figure 8.1). Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1223 construct.



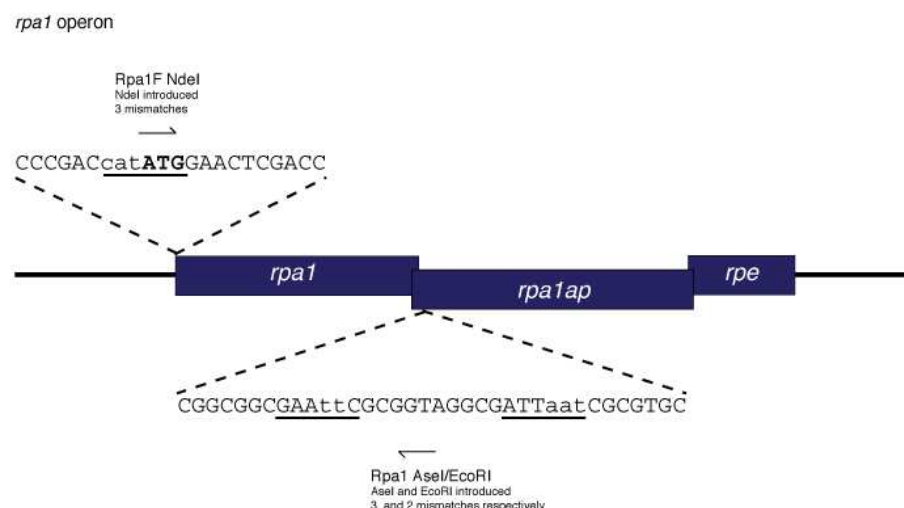
**Figure 8.9 Construction of the histidine-tagged *rpa1ap* plasmid**

A) Binding sites of the primers are shown B) the restriction sites in the overexpression vector used to insert amplified *rpa1ap* C) the resulting overexpression construct containing *rpa1ap* pTA1223. D) A basic sequence of construction.

### 8.1.2. Insertion of untagged *rpa1* upstream of histidine-tagged *rpa1ap*

#### Design of primers

The *rpa1* gene was amplified from the genomic *rpa1* operon clone pTA937 by touchdown PCR using the primers shown in Table 8.5, introducing a N-terminal *NdeI* and C-terminal *AseI* restriction sites. The restriction site *NdeI* was used since it contains the start codon ATG, an *NdeI* site is located in pTA963 (and pTA1223) directly upstream of the hexahistidine tag (see Figure 8.1).



**Figure 8.10 Introduction of *NdeI* and *AseI* sites**

The introduction of N-terminal *NdeI* and C-terminal *AseI* restriction sites maintains the reading frame of *rpa1*. Start codons shown in bold.

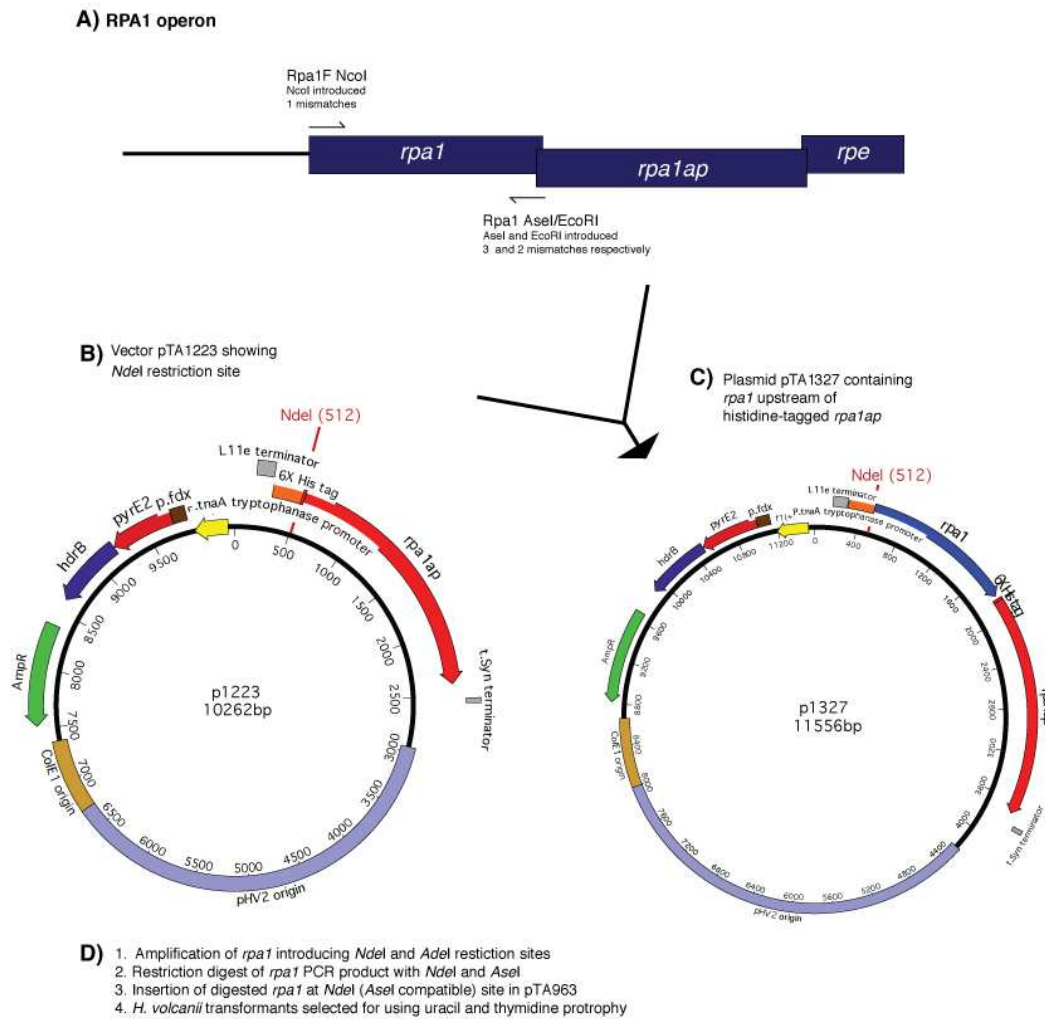
Primer	Sequence (5'-3')	Site inserted
Rpa1NdeI	CCCGACc <u>atATG</u> GAACTCGACC	<i>NdeI</i>
Rpa1 AseI/EcoR	CGGCGGCGA <u>Att</u> CGCGGTAGGCGAT <u>Taa</u> tCGCGTGC	<i>EcoRI/AseI</i>

**Table 8.5. Design of primers to insert *rpa1* into pTA1223**

The sequence for each primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of *rpa1* into pTA1223. Start codon of the *rpa1* is shown in bold.

#### **Insertion of untagged *rpa1* upstream of histidine-tagged *rpa1ap***

The PCR product of *rpa1*, using the primers in Figure 8.9 and Table 8.5, was gel purified, digested with *NdeI* and *AseI* and ligated into the plasmid pTA1223 using the *NdeI* site, and was transformed into *E. coli* XL-1 strain, as shown in Figure 8.11. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1327 construct.



**Figure 8.11 Construction of the untagged *rpa1* and histidine-tagged *rpa1ap* plasmid**

A) Binding sites of the primers are shown B) the restriction sites in the overexpression vector used to insert amplified *rpa1* C) the resulting overexpression construct containing untagged *rpa1* and histidine-tagged *rpa1ap* pTA1327. D) A basic sequence of construction.

### 8.3 Histidine-tag pulldowns of RPA1 and RPA1AP

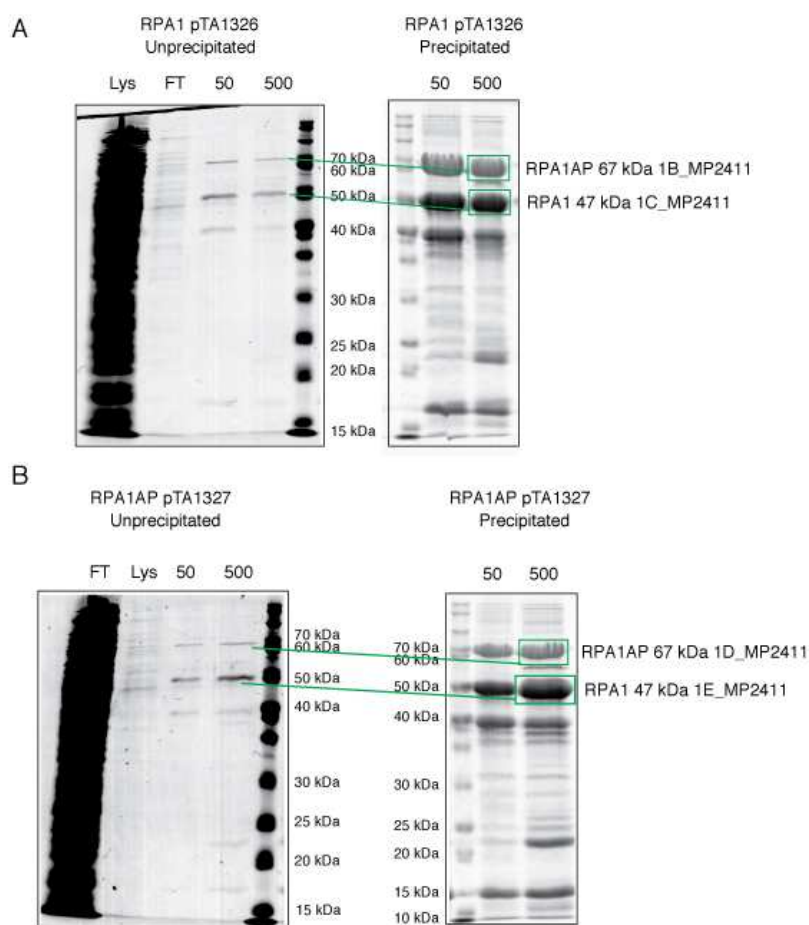
A starter culture was grown in Hv-Ca broth to an O.D of 0.1 and used to inoculate 333ml Hv-YPG broth, which was grown to an O.D of 0.4. This was followed by the addition of 25 mM tryptophan and a further incubation for 1 hour to induce expression of the histidine-tagged proteins. The culture was then centrifuged, the cell pellet was sonicated and the lysate incubated with nickel charged Sepharose beads. Thereafter the mixture was run through a gravity column, where the histidine-tagged proteins and

any interacting proteins remain bound to the nickel beads, thus allowing the purification of the histidine-tagged proteins and any interacting proteins from the cell lysate. After washing and elution of bound proteins using an imidazole step gradient, the purified sample was run on an SDS-PAGE gel. Precipitation using trichloroacetic acid and deoxycholate was used to enhance visualisation and identification of the eluted proteins.

Bands of interest were excised from the SDS-PAGE gel, reduced, carboxyamidomethylated and digested with Trypsin Gold. The resulting peptide sample was then analysed by mass spectrometry. This part of the work was performed by Susan Liddell (Proteomics Facility, School of Biosciences, University of Nottingham)

The histidine-tagged protein pulldowns (see Figure 8.12) showed that the histidine-tagged RPA1 interacts with untagged RPA1AP, this interaction is confirmed by the interaction between histidine-tagged RPA1AP and untagged RPA1. The other bands visible are also present in the empty vector control (see Figure 8.3) and consequently were ignored.





**Figure 8.12 Histidine-tagged RPA1 and RPA1AP pull-downs**

A) Histidine-tagged RPA1 pull down B) Histidine-tagged RPA1AP pull down. 50 and 500 mM imidazole were used for elution from the nickel column. Unprecipitated elutions are shown on the left, to enhance visualisation the protein elution were precipitated, shown on the right. Bands that were excised for mass spectrometry are indicated. See Appendix MP2411 for full mass spectrometry data, peptide matches for 1B-1E\_MP2511 are shown below.

Mass spectrometry was used to identify the bands as annotated in Figure 8.12, the matching peptides found are shown below. The results confirmed the interaction between RPA1 and RPA1AP, and vice versa. The interaction between RPA1 and RPA1AP must be stable *in vivo* and *in vitro*, since the untagged partner can be isolated without the aid of cross-linking agents.

## RPA1AP 1B\_MP2411

gene="rpalap"

locus\_tag="HVO\_1337"

Match to: [gi|292655491](#) Score: 671

**rpa-associated protein [Haloferax volcanii DS2]**

Found in search of 1B\_MP2411\_14SEPT11.pkl

Nominal mass ( $M_r$ ): **64829**; Calculated pI value: **3.49**

NCBI BLAST search of [gi|292655491](#) against nr

Sequence Coverage: **18%**

Matched peptides shown in **Bold Red**

1	MSASPVGTR	EIAYRVFAAE	<b>FDDASLSYSE</b>	<b>SDEERAPNYV</b>	<b>VTPTGAR</b> LNR
51	TFVAGVLTEV	EHVNDEVLRG	RIADPTGAFV	TYAGQYQPEP	MAYLDAATPP
101	AFVSLAGKAR	<b>TYEPDDADV</b>	<b>YSSVRPESVN</b>	<b>TVDADVRDRW</b>	<b>IVSAAEATLR</b>
151	<b>RIAVFDEALS</b>	<b>MPYRGDDLTR</b>	<b>ALEARGV DST</b>	<b>LAAGVPRAID</b>	<b>HYGTTR</b> TYLE
201	<b>ALREVAVQAL</b>	<b>ELVAGDR</b> DQV	DPLDVAPGDG	GDAVLGPLPE	LDLEPAASVD
251	IEVG EADAGE	VEADAGEFEN	ESGSEPAATL	DSEPEFDDGA	ESEAVEPESE
301	SEPIAEPDSE	PESASDPEPV	AAETTAAAES	QPGASESEPA	SEPSTTV DSE
351	PEPEPAAEPE	SEPAAEPDLT	SEPEPEP VTD	SLDTESEPDF	DDGALGDFED
401	DAPDAGATDA	GSLDDSGSGT	LGDFDDGFDD	PDPEAGDSTD	STDSASADAD
451	DSVDPDGMYE	LDEDERAEIE	SEFGTEFTSG	ADVDPAGEAD	IDVPDADDLT
501	EQLEDESAAA	AESDPEPATA	AAEPEPAVDS	EPEPTPDAAS	DGDEDEADAD
551	PDADESAEDI	DLESVVVDAM	DDLDDGDGAT	RDEVVAAVVD	EHGADPGAVE
601	DAIQEALLGG	RCYEPQDGV L	KAI		

## RPA1 1C\_MP2411

gene="rpal"

locus\_tag="HVO\_1338"

Match to: [gi|292655492](#) Score: 940

**replication protein A [Haloferax volcanii DS2]**

Found in search of 1C\_MP2411\_14SEPT11.pkl

Nominal mass ( $M_r$ ): **45954**; Calculated pI value: **4.20**

NCBI BLAST search of [gi|292655492](#) against nr

Sequence Coverage: **40%**

Matched peptides shown in **Bold Red**

1	MELDQHAEEL	ASALGV DKEE	<b>VKSDLQNLLQ</b>	<b>YSVPLDEAKQ</b>	SVRRKHGGGS
51	SGGSDGAPAT	KR <b>IVDIDPDG</b>	<b>GNVSVTVRVL</b>	TVGTR <b>SIVYQ</b>	<b>GDEQTIREGE</b>
101	LADESGVISY	TAWQDFGFEP	GDSVVIGNAG	VREWDGKPEL	NIGASSTVGV
151	ESETVETPYD	DR <b>IGGEADLI</b>	<b>DLQAGDRGRT</b>	VEVRVLEVDS	RTIDGR <b>NGET</b>
201	<b>TILSGVVADE</b>	<b>TGRLPFTDWA</b>	<b>PRPDVEEGAS</b>	<b>LRLSDVYVRE</b>	FR <b>GVPQVNLS</b>
251	<b>EFTTLDVLDD</b>	<b>PVSVTDSAPR</b>	LKIGEAVDAG	GMFDVELLGN	VLEV RDGSGL
301	IERCPKCSR V	VQNGQCR <b>QH G</b>	<b>EVDGEDDMRV</b>	<b>KAILDDGTGT</b>	<b>VTAILDRDLT</b>
351	EEIYGGTMAD	AMEAAREAMD	<b>KEVVADDIAD</b>	<b>TLVGREYRVR</b>	GNLSVDEYGA
401	NLETDEFEET	DDDP AER <b>AAA</b>	<b>LLTEVRA</b>		

## RPA1AP 1D\_MP2411

gene="rpa1ap"

locus\_tag="HVO\_1337"

Match to: [gi|292655491](#) Score: 578

**rpa-associated protein [Haloflex volcanii DS2]**

Found in search of 1D\_MP2411\_15SEPT11.pkl

Nominal mass (M<sub>r</sub>): **64829**; Calculated pI value: **3.49**

NCBI BLAST search of [gi|292655491](#) against nr

Sequence Coverage: **21%**

Matched peptides shown in **Bold Red**

```
1 MSASPVGTR EIAYRVFAAE FDDASLSYSE SDEERAPNYV VTPTGARLNR
51 TFVAGVLTEV EHVNDEVLRG RIADPTGAFV TYAGQYQPEP MAYLDAATPP
101 AFVSLAGKAR TYEPDDADV YSSVRPESVN TVDADVRDRW IVSAAEATLR
151 RIAVFDEALS MPYRGDDLTR ALEARGVDST LAAGVPRAID HYGTTTRTYLE
201 ALREVAVQAL ELVAGDRDQV DPLDVAPGDG GDAVLGPLE LDLEPAASVD
251 IEVGADAGE VEADAGEFEN ESGSEPAATL DSEPEFDDGA ESEAVEPESE
301 SEPIAEPDSE PESASDPEPV AAETTAAAES QPGASESEPA SEPSTTVDSSE
351 PEPEPAAEPE SEPAAEPDLT SEPEPEPVT DLDTESEPDF DDGALGDFED
401 DAPDAGATDA GSLDDSGSGT LGDFDDGFDD PDPEAGDSTD STDSASADAD
451 DSVDPDGMYE LDEDERAEIE SEFGTEFTSG ADVDPAGEAD IDVPDADDLT
501 EQLEDESAAA AESDPEPATA AAEPEPAVDS EPEPTPDAAS DGDEDEADAD
551 PDADESAEDI DLESVVVDAM DDLDDGDGAT RDEVVAHVVD EHGADPGAVE
601 DAIQEALLGG RCYEPQDGV L KAI
```

## RPA1 1E\_MP2411

gene="rpa1"

locus\_tag="HVO\_1338"

Match to: [gi|292655492](#) Score: 808

**replication protein A [Haloflex volcanii DS2]**

Found in search of 1E\_MP2411\_15SEPT11.pkl

Nominal mass (M<sub>r</sub>): **45954**; Calculated pI value: **4.20**

NCBI BLAST search of [gi|292655492](#) against nr

Sequence Coverage: **40%**

Matched peptides shown in **Bold Red**

```
1 MELDQHAEE L ASALGV DKEE VKSDLQNL LQ YSVPLDEAKQ SVRRKHGGGS
51 SGGSDGAPAT KRIVDIDPDG GNVSVTVRVL TVGTRSIVYQ GDEQTIREGE
101 LADESGVISY TAWQDFGFEP GDSVVIGNAG VREWDGKPEL NIGASSTVG
151 ESETVETPYD DRIGGEADLI DLQAGDRGRT VEVRVLEVDS RTIDGRNGET
201 TILSGVVADE TGRLPFTDWA PRPDVEEGAS LRLSDVYVRE FRGVPQVNLS
251 EFTTLDV LDD PVSVTDSAPR LKIGEAVDAG GMFDVELLGN VLEV RDGSG
301 IERCPKCSR V QNGQCRQHGE EVDGEDDMRV KAILDDGTGT VTAILDRDLT
351 EEIYGGTMAD AMEAAREAMD KEVVADDIAD TLVGREYRVR GNLSVDEYGA
401 NLETDEFEE T DDDPAERAAA LLTEVRA
```

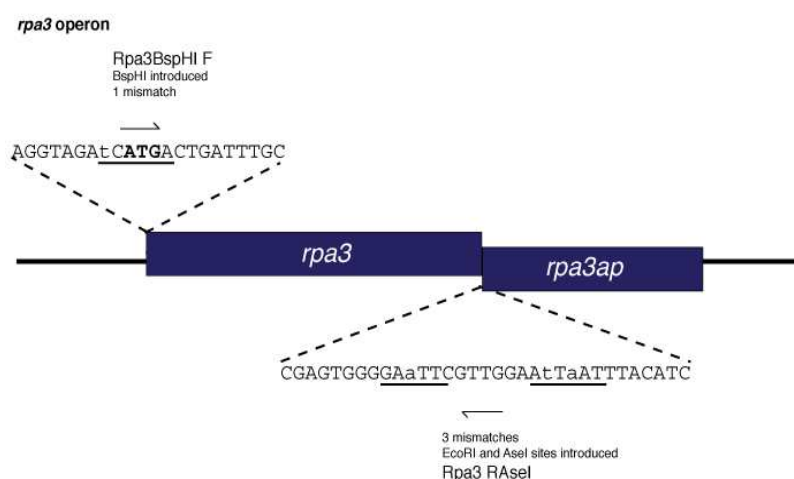
## 8.4. Construction of histidine-tagged *rpa3* with *rpa3ap*

To examine if RPA3 and RPA3AP interact with each other, an overexpression plasmid containing histidine-tagged *rpa3* and *rpa3ap* (in an operon) was constructed.

### 8.4.1. Construction of histidine-tagged *rpa3*

#### Design of primers

The *rpa3* gene was amplified from the genomic *rpa3* clone pTA884 by touchdown PCR using the primers shown in Table 8.6, introducing N-terminal *Nco*I and C-terminal *Eco*RI restriction sites. The restriction site *Bsp*HI was chosen, since it contains ATG followed by a G, to avoid changing the nucleotide G after the start codon (ATG) and altering the reading frame of the operon (see Figure 8.1).



**Figure 8.13**  
**Introduction of**  
***Bsp*HI and *Eco*RI**  
**sites**

The introduction of N-terminal *Bsp*HI and C-terminal *Eco*RI restriction sites maintain the reading frame of *rpa3*. Start codons shown in bold.

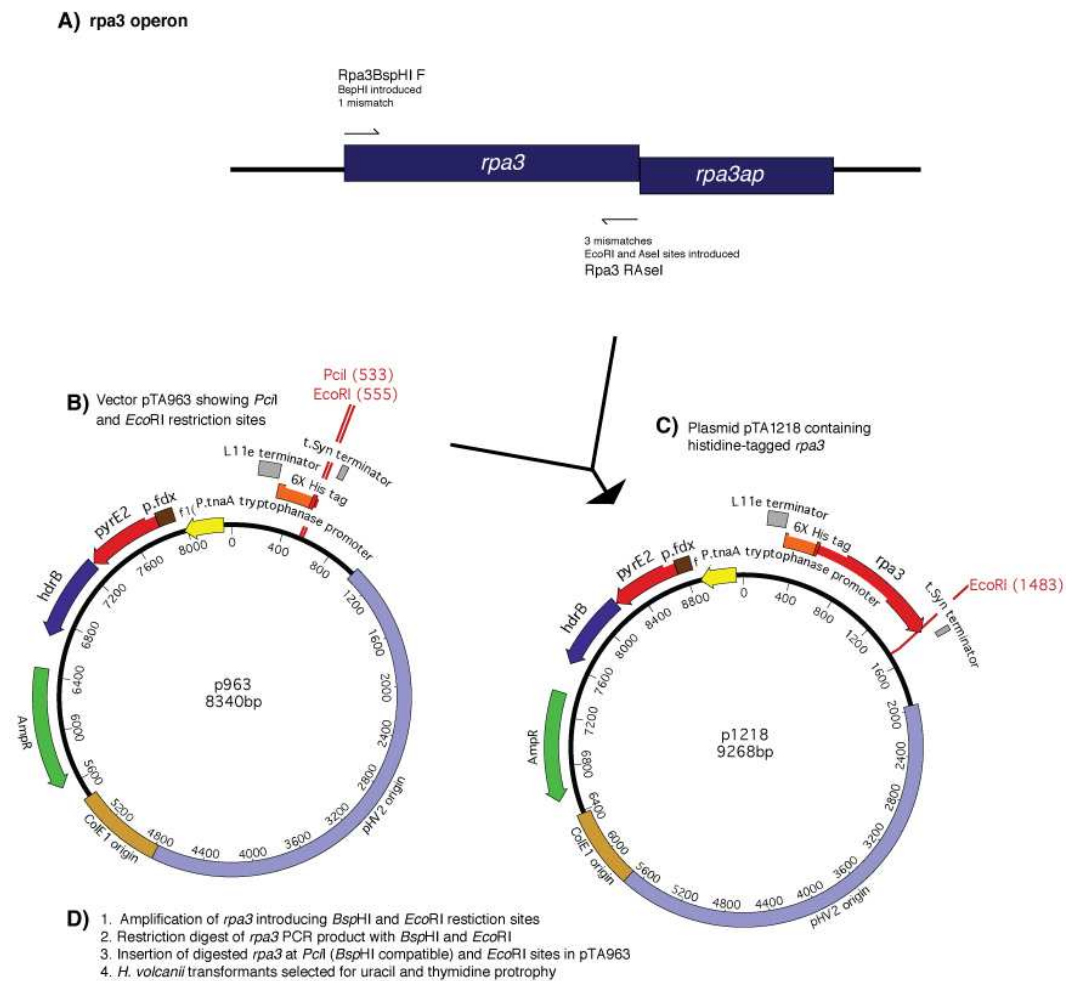
Primer	Sequence (5'-3')	Site inserted
Rpa3BspHI	AGGTAGAt <b>CATG</b> ACTGATTTC	<i>Bsp</i> HI
Rpa3 RAsel	CGAGTGGGGAaTTCGTTGGAAaTaATTACATC	<i>Ase</i> I

**Table 8.6. Design of primers to insert *rpa3* into pTA963**

The sequence for each primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of *rpa3* into pTA963. Start codon of the *rpa3* is shown in bold.

### Construction of histidine-tagged *rpa3* plasmid

The PCR product of *rpa3* was gel purified, digested with *Bsp*HI and *Eco*RI and ligated into the plasmid pTA963 using the *Pci*I and *Eco*RI sites, and was transformed into *E. coli* XL-1 strain, as shown in Figure 8.14 (see also Figure 8.1). Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1218 construct.



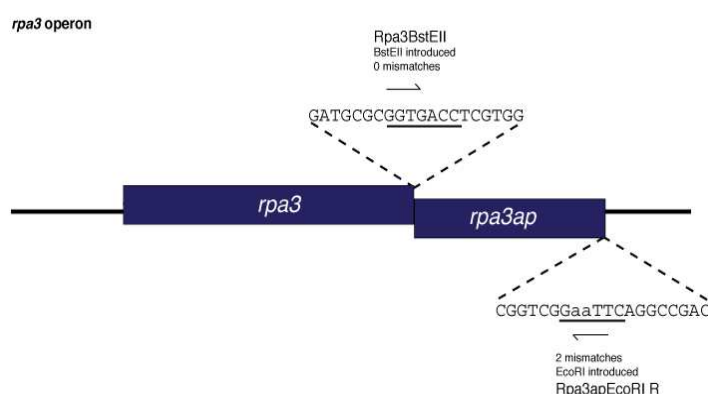
**Figure 8.14 Construction of the histidine-tagged *rpa3* plasmid**

A) Binding sites of the primers are shown, the introduction of *Ase*I is not relevant here and will be discussed later B) the restriction sites in the overexpression vector used to insert amplified *rpa3* C) the resulting overexpression construct containing *rpa3* pTA1218. D) A basic sequence of construction.

### 8.4.2. Insertion of untagged *rpa3ap* downstream of histidine-tagged *rpa3*

#### Design of primers

The *rpa3ap* gene was amplified from the genomic *rpa3* operon clone pTA884 by touchdown PCR using the primers shown in Table 8.7, featuring N-terminal *Bst*EII and C-terminal *Eco*RI restriction sites. The naturally occurring restriction site *Bst*EII was chosen since it does not involve any base-pair changes and maintains the reading frame of the *rpa3* operon.



**Figure 8.15 Introduction of internal *Bst*EII and *Bam*HI sites**

The use of N-terminal *Bst*EII and C-terminals *Bam*HI restriction sites maintains the reading frame of *rpa3ap*. Start codons shown in bold.

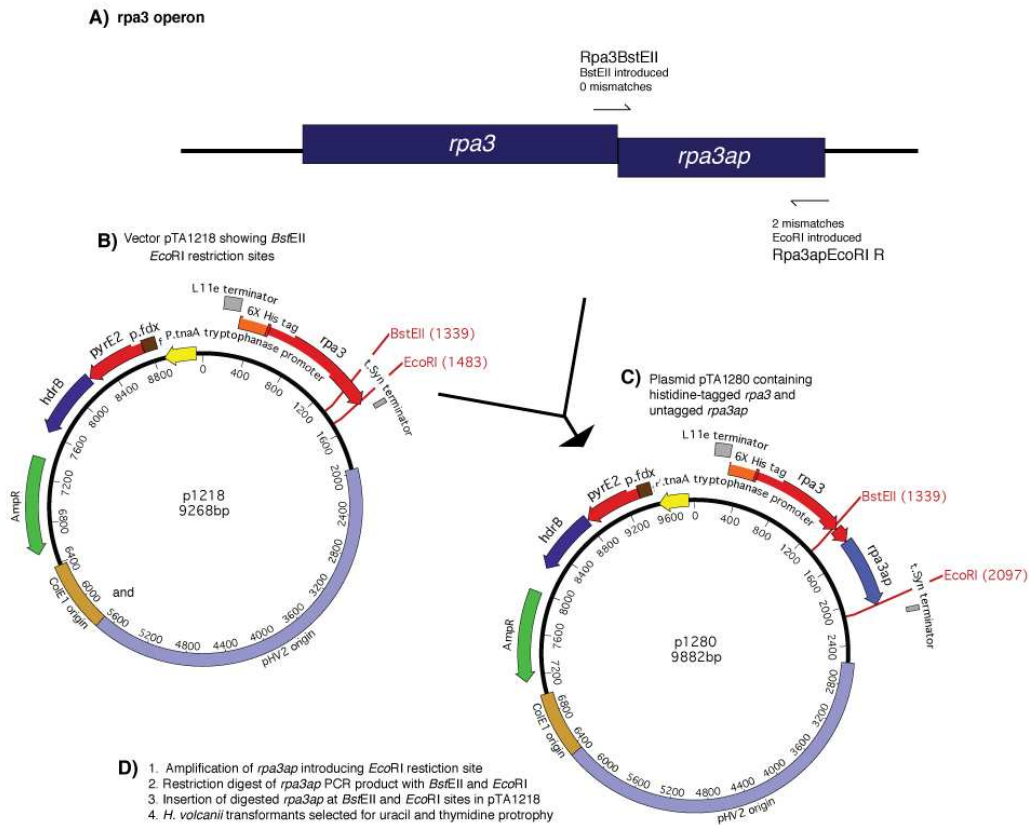
Primer	Sequence (5'-3')	Site inserted
Rpa3BstEII	GATGCGCGGTGACCTCGTGG	<i>Bst</i> EII
Rpa3apEcoRI R	CGGTCGGaaTTCAGGCCGAC	<i>Eco</i> RI

**Table 8.7 Design of primers to insert *rpa3ap* into pTA1280**

The sequence for each primer is shown with the base changes made to introduce restriction sites, (in lower case), restriction site, (underlined), to allow cloning of *rpa3ap* into pTA1280.

#### Insertion of *rpa3ap* downstream of histidine-tagged *rpa3*

The PCR product of *rpa3ap* was gel purified, digested with *Bst*EII and *Eco*RI and ligated into the plasmid pTA1218 using the *Bst*EII and *Eco*RI sites, and was transformed into *E. coli* XL-1 strain, as shown in Figure 8.16. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1280 construct.



**Figure 8.16 Insertion of untagged *rpa3ap* into pTA1218**

A) Binding sites of the primers are shown B) the restriction sites in the overexpression vector used to insert amplified *rpa3ap* C) the resulting overexpression construct containing histidine-tagged *rpa3* and untagged *rpa3ap* pTA1280. D) A basic sequence of construction.

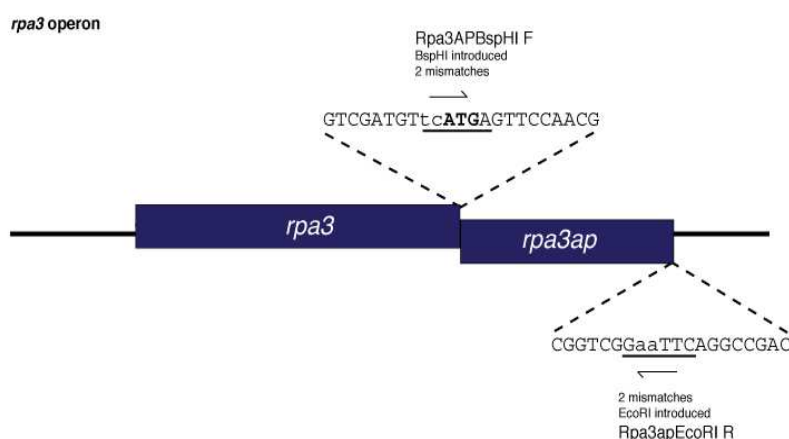
## 8.5. Construction of *rpa3* with histidine-tagged *rpa3ap*

To examine RPA3AP interacts with RPA3, an overexpression plasmid containing *rpa3* and histidine-tagged *rpa3ap* (in an operon) was constructed.

### 8.5.1. Construction of histidine-tagged *rpa3ap*

#### Design of primers

The *rpa3ap* gene was amplified from the genomic *rpa3* operon clone pTA884 by touchdown PCR using the primers shown in Table 8.8, introducing N-terminal *Bsp*HI and C-terminal *Eco*RI restriction sites. The restriction site *Bsp*HI was chosen, since it contains ATG followed by an A, to avoid changing the nucleotide A after the start codon (ATG) and altering the reading frame of the operon (see Figure 8.1).



**Figure 8.17**  
**Introduction of**  
***Bsp*HI and *Eco*RI**  
**sites**

The introduction of N-terminal *Bsp*HI and C-terminal *Eco*RI restriction sites maintains the reading frame of *rpa3ap*. Start codons shown in bold.

Primer	Sequence (5'-3')	Site inserted
Rpa3APBspHI F	GTCGATGTt c <b>ATG</b> AGTTCCAACG	<i>Bsp</i> HI
Rpa3apEcoRI R	CGGTCGGaa <b>TTC</b> AGGCCGAC	<i>Eco</i> RI

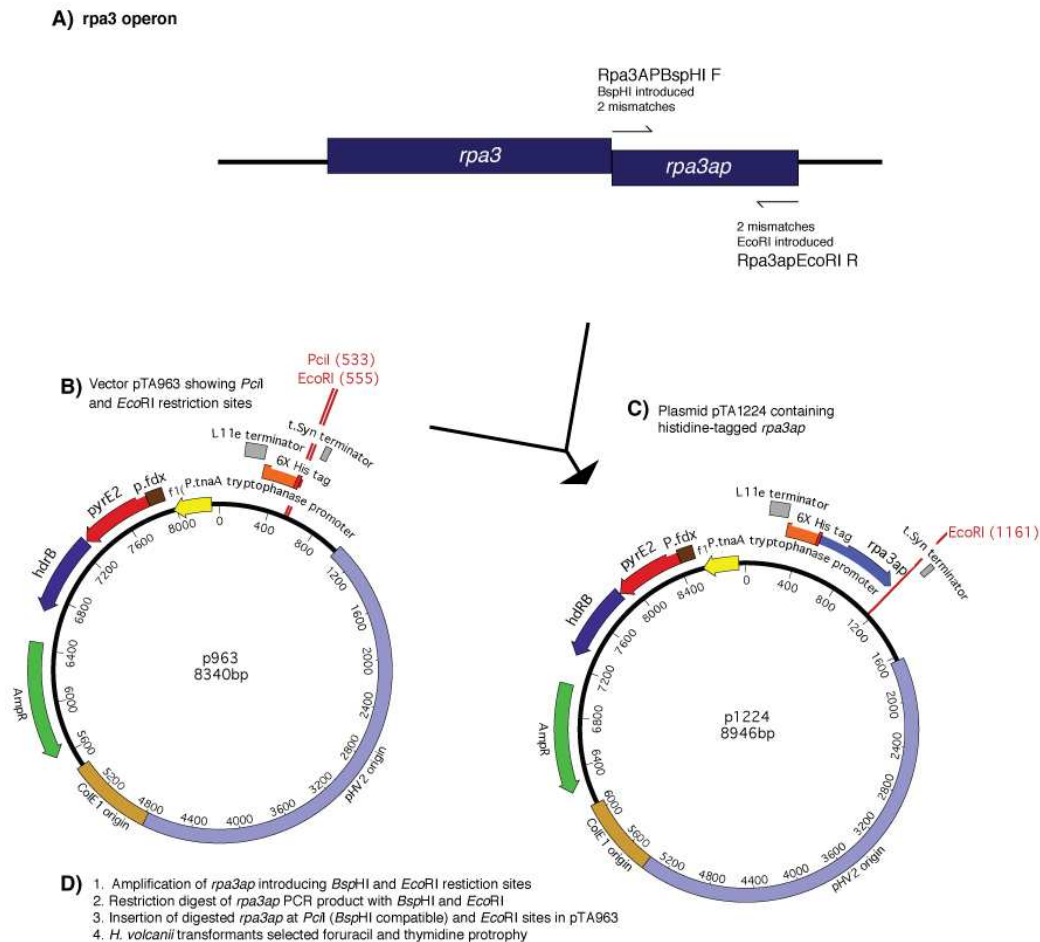
**Table 8.8 Design of primers to insert *rpa3ap* into pTA963**

The sequence for each primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of *rpa3ap* into pTA963. Start codon of the *rpa3ap* is shown in bold.

### Construction of histidine-tagged *rpa3ap* plasmid

The PCR product of *rpa3ap* was gel purified, digested with *Bsp*HI and *Eco*RI, and ligated into the plasmid pTA963 using the *Pci*I and *Eco*RI sites, followed by transformation into *E. coli* XL-1 strain, as shown in Figure 8.18 (see also Figure 8.1). Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1224 construct.





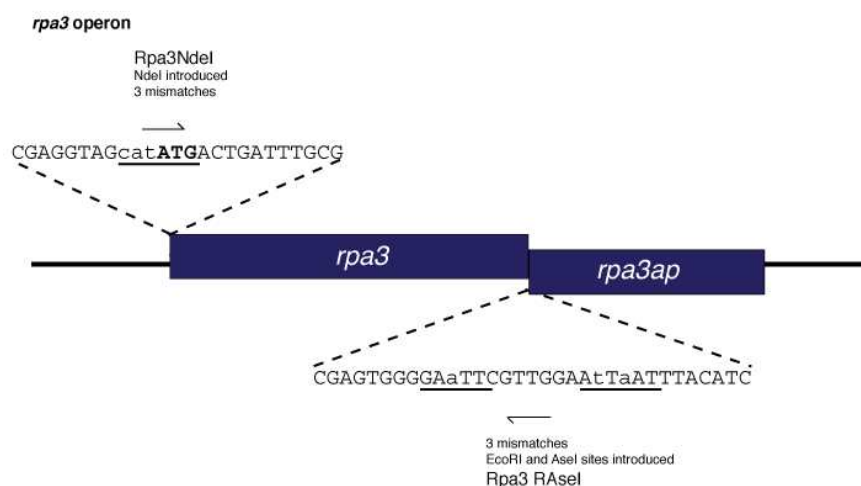
**Figure 8.18 Construction of the histidine-tagged *rpa3ap* plasmid**

A) Binding sites of the primers are shown B) the restriction sites in the overexpression vector used to insert amplified *rpa3ap* C) the resulting overexpression construct containing *rpa3ap* pTA1224. D) A basic sequence of construction.

### 8.5.2. Insertion of untagged *rpa3* upstream of histidine-tagged *rpa3ap*

#### Design of primers

The *rpa3* gene was amplified from the genomic *rpa3* operon clone pTA884 by touchdown PCR using the primers shown in Table 8.9, introducing N-terminal *NdeI* and C-terminal *AseI* restriction sites. The restriction site *NdeI* was used since it contains the start codon ATG and is the site directly upstream of the histidine tag (see Figure 8.1), it is compatible with the *AseI* site.



**Figure 8.19 Introduction of *NdeI* and *AseI* sites**

The introduction of N-terminus *NdeI* and C-terminus *AseI* restriction sites maintains the reading frame of *rpa3*. Start codons shown in bold.

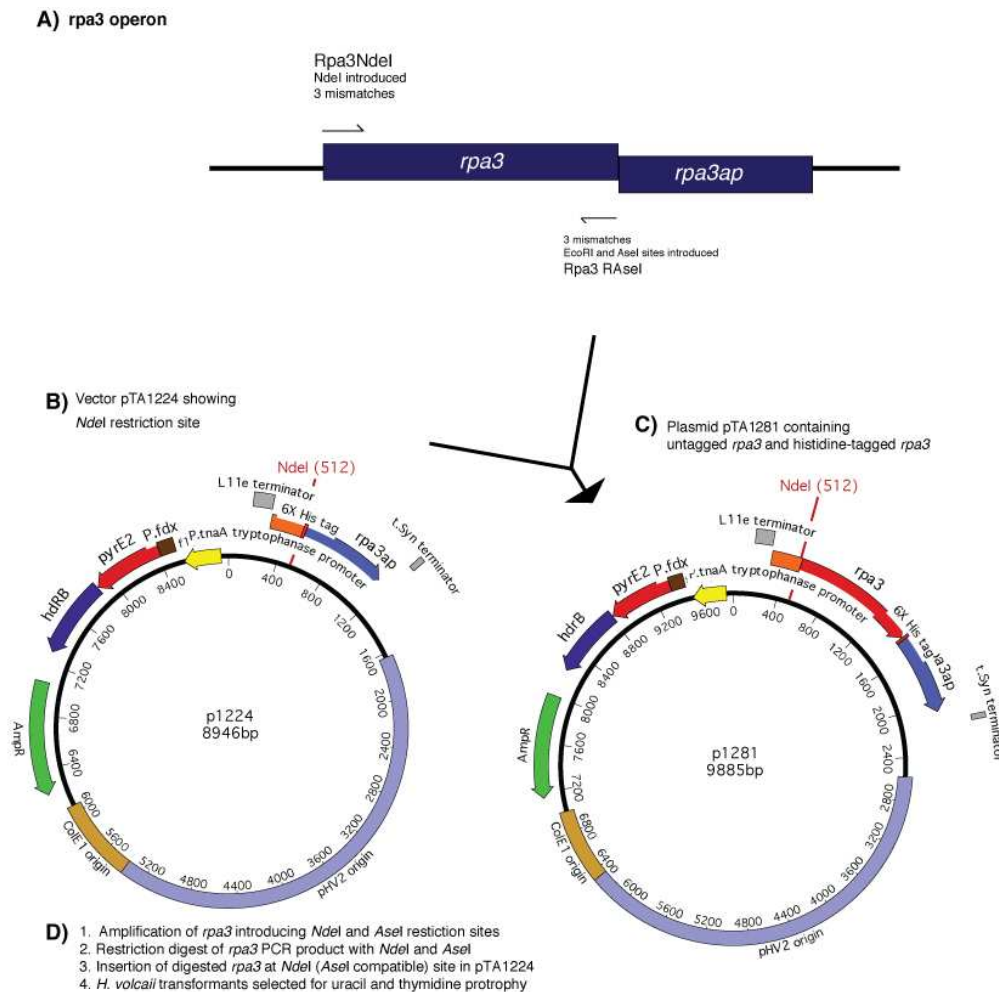
Primer	Sequence (5'-3')	Site inserted
Rpa3NdeI	CGAGGTAGcat <b>ATG</b> ACTGATTTGCG	<i>NdeI</i>
Rpa3 RAsel	CGAGTGGGGAaTTCGTTGGAAtTaATT ACATC	<i>EcoRI/AseI</i>

**Table 8.9 Design of primers to insert *rpa3* into pTA1281**

The sequence for each primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of *rpa3* into pTA1281. Start codon of the *rpa3* is shown in bold.

#### **Insertion of untagged *rpa3* upstream of histidine-tagged *rpa3ap***

The PCR product of *rpa3* was gel purified, digested with *NdeI* and *AseI* and ligated into the plasmid pTA1224 using the *NdeI* site, and was transformed into *E. coli* XL-1 strain, as shown in Figure 8.20. Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA1281 construct.



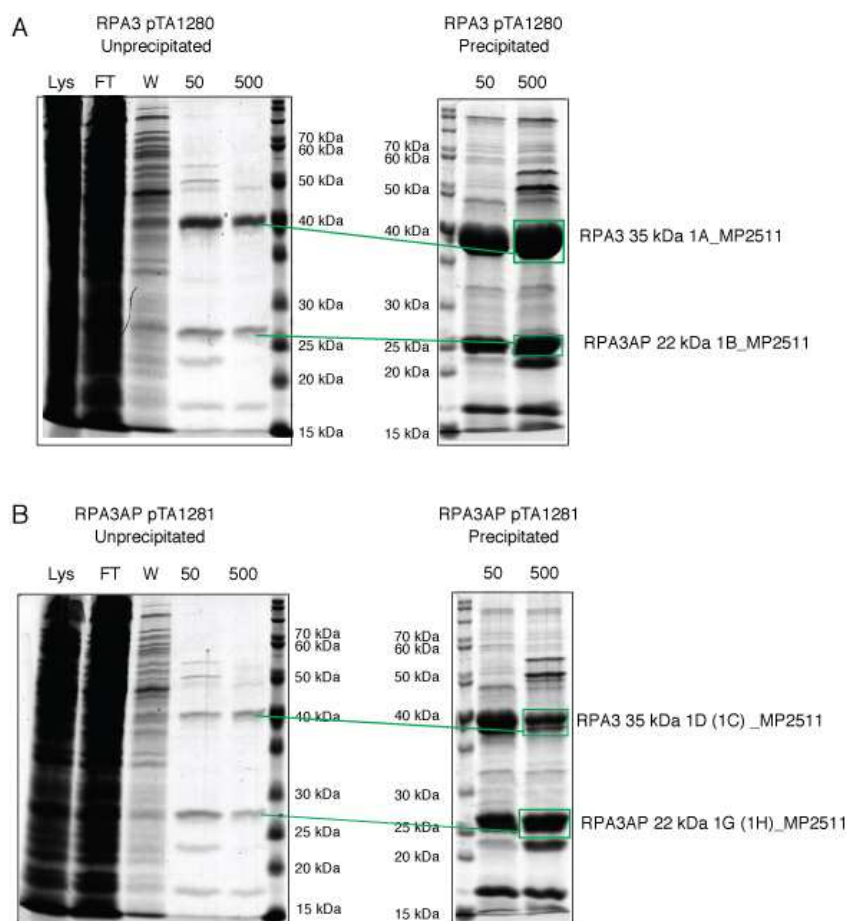
**Figure 8.20 Construction of untagged *rpa3* and histidine-tagged *rpa3ap* plasmid**

A) Binding sites of the primers are shown B) the restriction sites in the overexpression vector used to insert amplified *rpa3* C) the resulting overexpression construct containing untagged *rpa3* and histidine-tagged *rpa3ap* pTA1281. D) A basic sequence of construction.

## 8.6 Histidine-tag pulldown of RPA3 and RPA3AP

The histidine-tagged RPA3 and RPA3AP pull-downs (see Figure 8.21) were performed as described for RPA1 and RPA1AP histidine-tagged protein pull-downs. The pulldowns showed that the histidine-tagged RPA3 interacts with the untagged RPA3AP, this interaction was confirmed by the interaction between the histidine-tagged RPA3AP and untagged RPA3. In addition the interaction between RPA1 and RPA1AP must be stable *in vivo* and *in vitro*, since the untagged partner can be isolated without the aid of cross-linking agents. The other bands visible, including 1C,

are also present in the empty vector control (see Figure 8.3) and consequently were ignored. A band (1H) of slightly greater mobility than the RPA3AP band (1G) is assumed to be a breakdown product resulting from the histidine-tagged RPA3AP.



**Figure 8.21 Histidine-tagged RPA3 and RPA3AP pull-downs**

A) Histidine-tagged RPA3 pull down B) Histidine-tagged RPA3AP pull down. 50 and 500 mM imidazole were used for elution from the nickel column. Unprecipitated elutions are shown on the left, to enhance visualisation the protein elution were precipitated, shown on the right. Bands that were excised for mass spectrometry are indicated. See Appendix MP2511 for full mass spectrometry data.

## RPA3 1A\_MP2511

gene="rpa3"

locus\_tag="HVO\_0292"

Match to: [gi|292654472](#) Score: 796

**replication protein A [Haloferax volcanii DS2]**

Found in search of 1A\_MP2511\_26SEPT11.pkl

Nominal mass ( $M_r$ ): **34562**; Calculated pI value: **4.20**

NCBI BLAST search of [gi|292654472](#) against nr

Sequence Coverage: **43%**

Matched peptides shown in **Bold Red**

```
1 MTDLRTHAAE IADQFSDHLD VSADEVEERL ESLVTEYRVP VDEARRSVVN
51 SYLDEAGIER DELAGGSGGN EQTLLNDIDE DEQWVDVRK VVELWEPRSE
101 SIAQVGLLGD DSGRMKFVSF TTSELPELEE GKSYALGNVV TDEYQGNFSV
151 KLNRTTSITE LDEEIEVGDD STSVEGALVD SQSGSGLIKR CPEEGCTRVL
201 QNGRCSEHGS VEGEFDLRIK AVVDDGDEVH EVIFNREMTE ELTGIELDEA
251 KQMAMDALDT TIVEEEMRGD LVGYYYRVTG PTLGRYVLN EVERLREPAD
301 AEELLIKARS M
```

## RPA3AP 1B\_MP2511

gene="rpa3ap"

locus\_tag="HVO\_0291"

Match to: [gi|292654471](#) Score: 484

**rpa-associated protein [Haloferax volcanii DS2]**

Found in search of 1B\_MP2511\_26SEPT11.pkl

Nominal mass ( $M_r$ ): **21979**; Calculated pI value: **4.30**

NCBI BLAST search of [gi|292654471](#) against nr

Sequence Coverage: **59%**

Matched peptides shown in **Bold Red**

```
1 MSSNEIPTRE VARRVFAQEF NDAGYTFKES DDERAPVYLL LPTGESANRV
51 FLVGTLTEKE DVGEDNEYWR GRIVDPTGTF FVYAGQYQPE AASALRDLDA
101 PAYVAVVGKP RTYETDDGSI NVSVRPESIT EVDAASTRDW VTETADKTL
151 RIAAFDEEGD EYARMAREHY DLDPEEYKRA AIAALESLEQ ADELSA
```

## RPA3 1D\_MP2511

gene="rpa3"

locus\_tag="HVO\_0292"

Match to: [gi|292654472](#) Score: 293

**replication protein A [Haloferax volcanii DS2]**

Found in search of 1D\_MP2511\_26SEPT11.pkl

Nominal mass ( $M_r$ ): **34562**; Calculated pI value: **4.20**

NCBI BLAST search of [gi|292654472](#) against nr

Sequence Coverage: **23%**

Matched peptides shown in **Bold Red**

```
1 MTDLRTHAAE IADQFSDHLD VSADEVEERL ESLVTEYRVP VDEARRSVVN
51 SYLDEAGIER DELAGGSGGN EQTLLNDIDE DEQWVDVRK VVELWEPRSE
101 SIAQVGLLGD DSGRMKFVSF TTSELPELEE GKSYALGNVV TDEYQGNFSV
151 KLNRTTSITE LDEEIEVGDD STSVEGALVD SQSGSGLIKR CPEEGCTRVL
201 QNGRCSEHGS VEGEFDLRIK AVVDDGDEVH EVIFNREMTE ELTGIELDEA
251 KQMAMDALDT TIVEEEMRGD LVGYYYRVTG PTLGRYVLN EVERLREPAD
301 AEELLIKARS M
```

## RPA3AP 1G\_MP2511

gene="rpa3ap"

locus\_tag="HVO\_0291"

Match to: [gi|292654471](#) Score: 435

**rpa-associated protein [Haloferax volcanii DS2]**

Found in search of 1G\_MP2511\_26SEPT11.pkl

Nominal mass ( $M_r$ ): **21979**; Calculated pI value: **4.30**

NCBI BLAST search of [gi|292654471](#) against nr

Sequence Coverage: **53%**

Matched peptides shown in **Bold Red**

```
1  MSSNEIPTRE VARRVFAQEF NDAGYTFKES DDERAPVYLL LPTGESANRV
51 FLVGTLTEKE DVGEDNEYWR GRIVDPTGTF FVYAGQYQPE AASALRDLDA
101 PAYVAVVGKP RTYETDDGSI NVSVRPESIT EVDAATRDRW VTETADKTLD
151 RIAAFDEEGD EYARMAREHY DLDP E EYKRA AIAALESLEQ ADELSA
```

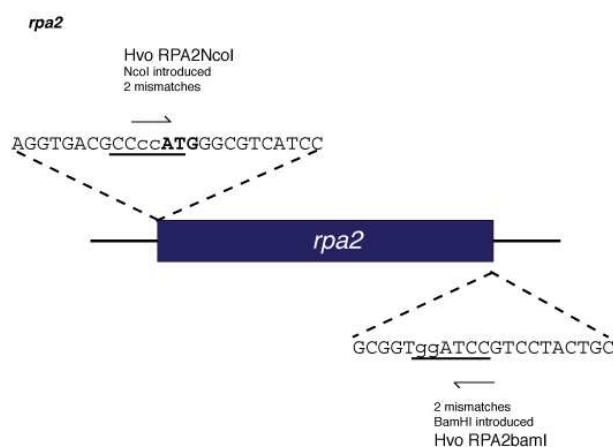
## 8.7. Construction of histidine-tagged *rpa2* plasmid

To examine if RPA2 interacts with RPA1, RPA3 and/or their associated proteins, as well as to identify protein interactions that might indicate a cellular role for RPA2, a histidine-tagged RPA2 construct was generated.

### 8.7.1. Construction of histidine-tagged *rpa2*

#### Design of primers

The *rpa2* gene was amplified from the genomic *rpa2* clone pTA889 by touchdown PCR using the primers shown in Table 8.10, allowing the introduction of a N-terminal *Nco*I and C-terminal *Bam*HI restriction sites. The restriction site *Nco*I was chosen, since it contains ATG followed by a G, to avoid changing the nucleotide G after the start codon (ATG).



**Figure 8.22 Introduction of *Nco*I and *Bam*HI sites**

The introduction of N-terminal *Nco*I and C-terminal *Bam*HI restriction sites maintains the reading frame of *rpa2*. Start codons shown in bold.

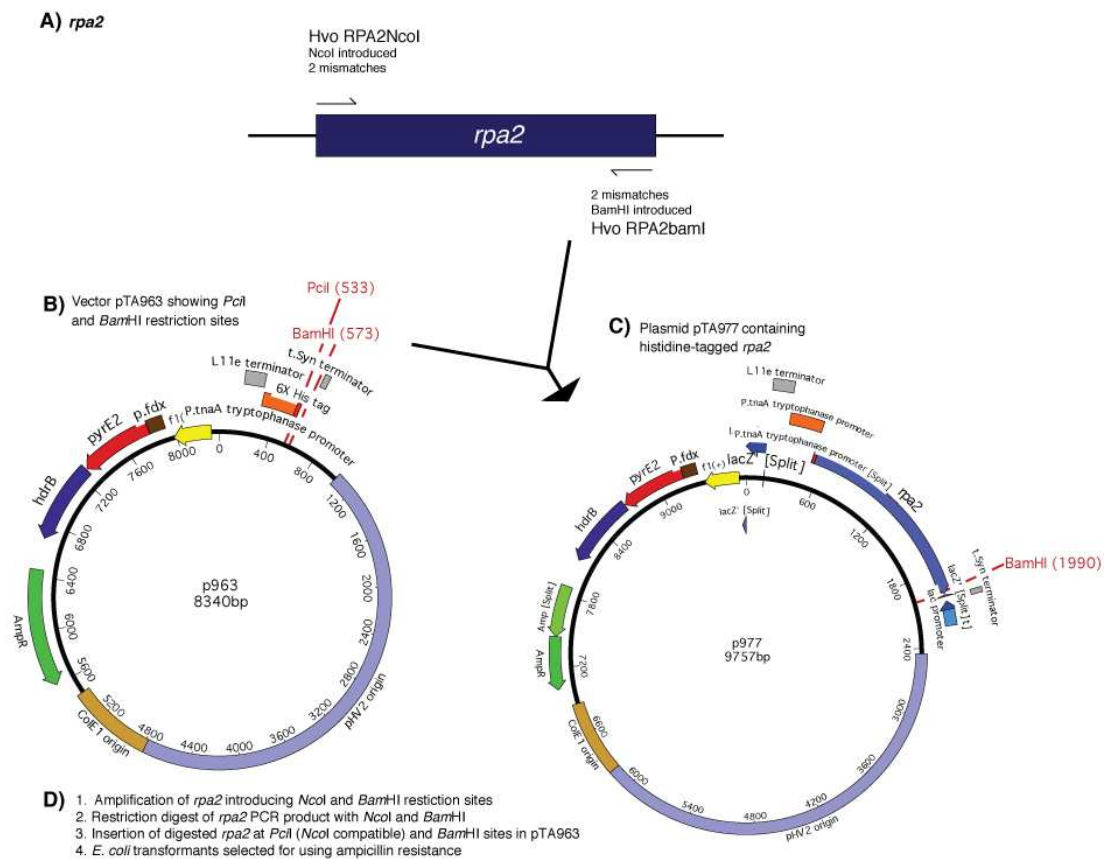
Primer	Sequence (5'-3')	Site inserted
Hvo RPA2NcoI	AGGTGACGCCcc <b>AT</b> GGGCGTCATCC	<i>Nco</i> I
Hvo RPA2bamI	GCGGTgg <b>AT</b> CCGTCCTACTGC	<i>Bam</i> HI

**Table 8.10 Design of primers to insert *rpa2* into pTA963**

The sequence for each primer is shown with the base changes, (in lower case), made to introduce a novel restriction site, (underlined), to allow cloning of *rpa2* into pTA963. Start codon of the *rpa2* is shown in bold.

## Construction of histidine-tagged *rpa2* plasmid

The PCR product of *rpa2* was gel purified, digested with *Nco*I and *Bam*HI, and ligated into the plasmid pTA963 using the *Pci*I and *Bam*HI sites, followed by transformation into *E. coli* XL-1 strain, as shown in Figure 8.23 (see also Figure 8.1). Restriction digest of plasmid extractions from transformants followed by sequencing was used to screen for the correct pTA977 construct.



**Figure 8.23 Construction of the histidine-tagged *rpa2* plasmid**

A) Binding sites of the primers are shown B) the restriction sites in the overexpression vector used to insert amplified *rpa2* C) the resulting overexpression construct containing *rpa2* pTA977. D) A basic sequence of construction.

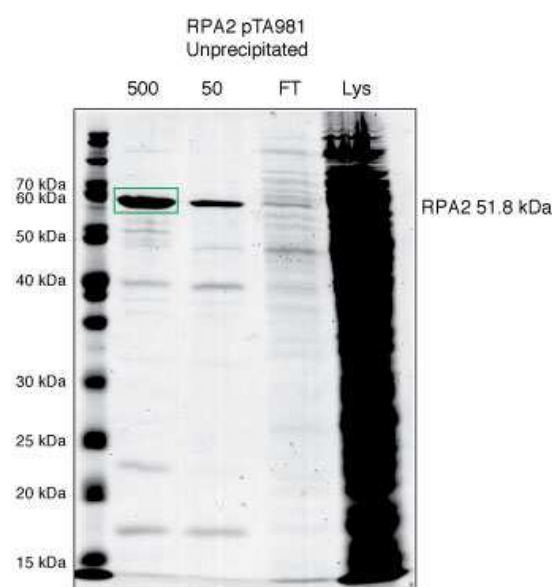
## 8.8 Histidine-tag pulldown of RPA2

The histidine-tag pull-down was performed as described for RPA1 and RPA1AP.

Histidine-tag pulldowns shows that RPA2 (see Figure 8.24) does not interact strongly with RPA1, RPA3 or any of the associated proteins, since none of these could be



identified by mass spectrometry. The other bands visible are also present in the empty vector control (see Figure 8.3) and consequently were ignored.



**Figure 8.24 Histidine-tag pull downs of RPA2**

Histidine-tagged RPA2 pulldown, see below for mass spectrometry data on RPA2 band 1A\_MP2111. 50 and 500 mM imidazole were used for elution off the nickel column. Band that was excised for mass spectrometry is outlined.

The mass spectrometry data shown below confirms that the band annotated on the gel in RPA2, it was not necessary to precipitate the elution samples.

### RPA2 1A\_MP2111

gene="rpa2"

locus\_tag="HVO\_0519"

Match to: [gi|292654691](#) Score: 727

**replication protein A [Haloflex volcanii DS2]**

Found in search of 1A\_MP2411\_14SEPT11.pkl

Nominal mass ( $M_r$ ): **51007**; Calculated pI value: **3.96**

NCBI BLAST search of [gi|292654691](#) against nr

Sequence Coverage: **34%**

Matched peptides shown in **Bold Red**

```

1  MGVIREFVYDD  LDTDFEFEEF  EAAVNDKVEQ  MGGLADEETA  AMLIAHELRD
51  EEVNGIADIE  PGMEDVKFLA  KVSIGEVRT  FERDGEDEDG  RVVNIEVADE
101 TGRIRVSLWD  EMAAGAKENL  EVGTVLRIGG  RPKDGYNGVE  VSASKVEEDL
151  DAEVDVQVLD  SYRVEDLALG  LSDVNLEGKI  LDAGTVRTFD  RDDGTEGRVS
201 NLSVGDPTGR  VRVTLWDERA  DLAEELEPGQ  SVEVVDGYVR  ERDGSLELHV
251 GSRGAVEPID  EDIEYVPETT  DIGSLELGQT  VDIAGGVIEA  DGKRTFDRDD
301 GSEGQVRNIR  VKDGTGDIRV  ALWGEKADAD  VDLADYVVIT  DAEIKEGWQE
351  DLEASAGWRS  SVAVMDEAPE  GAAGTDAGGS  APTPPSDEGL  GAFSGDGSSD
401  DTSAANGGSS  SDASAAESTG  EAVEFTGTVV  QAGTPVILDD  GTQTKTVDTD
451 ADLGLGDEVT  VSGTETDGRI  SAESVEVHTG  AQQ

```

## 8.9. Discussion

The hexahistidine-tagged protein pull-downs demonstrate clearly that RPA1 interacts with RPA1AP and vice versa, and that RPA3 interacts with RPA3AP and vice versa, supporting the hypothesis that RPA1 and RPA3 form complexes with their respective associated proteins (APs). This novel interaction between RPA and their associated proteins is the first to be discovered in archaea, and provides an insight into how other RPA and APs might function in archaea. Furthermore, this finding indicates that the heterotrimeric RPA found in *Pyrococcus furiosus* might be an exception amongst archaea (Komori and Ishino 2001).

However, the interaction of the histidine-tagged RPA3AP with RPA3 is not as strong as that of histidine-tagged RPA3 with RPA3AP. This may be due to perturbation of the RPA3AP structure owing to the presence of the N-terminal hexahistidine-tag, which could hinder the RPA3AP interaction with RPA3. This is supported by the presence of an RPA3AP breakdown product in the histidine-tagged RPA3AP pull-down. A C-terminal histidine-tagged RPA3AP could be used instead, as this may interfere less with RPA3AP structure and consequently will not constrain the RPA3AP:RPA3 interaction.

The hexahistidine-tagged pulldowns of RPA1, RPA2 and RPA3 also reinforce the hypothesis that they do not form a RPA heterotrimer as seen in eukaryotes and *P. furiosus*, since none of the three *H. volcanii* RPAs co-purify in the pull-downs. The failure of the hexahistidine-tagged RPA2 to pull down any interacting proteins prevents further speculation about specific cellular roles. However, the absence of RPA1 and 3 (as well as RPA1AP and RPA3AP) in the RPA2 protein pulldowns supports the phylogenetic analysis in Chapter 3, indicating an evolutionary divergence of RPA2 from RPA1 and 3 for the adaption of different cellular functions.

## Discussion

The aims of this study were to determine whether the three RPAs of *H. volcanii* form a heterotrimeric complex as in *P. furiosus*, or alternatively whether each RPA acts on its own; in the latter case, RPA1 and RPA3 might form a complex with their respective associated proteins. This included examining if all three of the RPAs are essential. In particular it was examined whether RPA1 and 3 are redundant with each other, the latter was a strong possibility since these two RPAs are closely related. Furthermore, what is the role of associated proteins (APs)? Are they required for their respective RPA to function? Lastly, what is the function of the RPE and why is it included in the *rpa1* operon?

The ability to delete both the *rpa1* and *rpa3* operons with relative ease, but not *rpa2*, indicates that the cell's requirement for each RPA is not equal, making it highly unlikely that they function collectively. The complete inability to delete *rpa2* suggests that RPA2 is critical for cell survival and most probably plays a significant role in DNA replication. It appears that the level of RPA2 in the cell is tightly regulated, as it was not possible to maintain an additional copy of *rpa2* on the chromosome. For example, too little RPA2 might hinder DNA replication since ssDNA would be unstable and susceptible to degradation by nucleases, whereas too much RPA2 would outcompete DNA replication proteins in ssDNA-binding. This would explain why the insertion of *rpa2* at the ectopic *pyrE2* locus was unsuccessful, since an additional copy of *rpa2* on the chromosome would have resulted in an elevation of *rpa2* expression. The newly devised method shown in Figures 6.28 and 6.29 should circumvent this problem, since a copy of *rpa2* under control of the tryptophanase promoter would be used to replace *rpa2* at its native locus, using positive selection.

The ease at which the *rpa1*, *rpa1ap*, *rpe* and the *rpa1* operon deletion mutants were made, coupled with a lack DNA damage sensitivity, signifies that the *rpa1* operon does not play a major role in DNA replication or repair. The slight growth defect of the *rpa1* operon deletion mutant demonstrates that the products of the *rpa1* operon aid cell growth and duplication but are not essential. Similarly, the ability to generate *rpa3*, *rpa3ap* and the *rpa3* operon deletion mutants indicates that the *rpa3* operon is not essential for cell survival. However the significant sensitivity to DNA damage

shown by the individual *rpa3*, *rpa3ap* and the *rpa3* operon mutants indicates the *rpa3* operon plays a significant role in DNA repair. Furthermore, both the individual  $\Delta rpa3$  and  $\Delta rpa3ap$  mutants show a similar sensitivity to DNA damage to each other and to the *rpa3* operon mutant, providing genetic evidence that RPA3 and RPA3AP act in the same DNA repair pathway. The slight growth defect exhibited by these mutants shows the extension of this role to cell growth and duplication.

The generation of the double mutant featuring a deletion of both the *rpa1* operon and the *rpa3* operon proved to be difficult, and the presence of the complementary *rpa1* operon episomal plasmid was essential to generate this double mutant. However, the sensitivity to UV and MMC DNA damage shown by the double operon deletion mutant was no greater than that shown by the *rpa3* operon mutant. This suggests that there is an element of redundancy between the two operons, and that they might function in the same pathway of DNA repair. Therefore, the results of this study have established that the products of the *rpa1* and the *rpa3* operon have overlapping roles in DNA repair; if they had differing roles, the double *rpa1 rpa3* operon deletion mutant would have been more sensitive to DNA damage than either single mutant. However, the individual operon deletions indicate the products of the *rpa3* operon function predominantly in DNA repair.

The inability to make *rpa1 rpa3ap* and *rpa1ap rpa3* deletion mutants suggests that the APs can only function with their respective RPA. Had it been possible to make these deletion mutants, it would have indicated that the APs are not RPA-specific, and that the *rpa1* and *rpa3* operon are fully redundant. Furthermore, the specific interaction between the RPAs and their respective APs is supported biochemically. Both the histidine-tagged RPA1 and RPA3 co-purified with their respective untagged associated proteins (AP), and similarly the histidine-tagged APs co-purified with their respective untagged RPAs. This finding supports the hypothesis that in *H. volcanii*, RPA1 and RPA3 form complexes with their respective APs. Therefore, this study has shown genetically and biochemically that the APs interact with the RPAs, and that this interaction is RPA-specific. This is the first report investigating the function of the archaeal COG3390 associated proteins (APs), thus providing an important insight of the structure and function of *H. volcanii* single-strand DNA binding proteins.

## Future Perspectives

The characterisation in this study of the three RPAs of *H. volcanii* and their associated proteins has laid the foundations for further investigations. The identification of a physical interaction between RPAs and APs is the first of its kind to be discovered in archaea. However, at this point it is not known whether the interaction between the RPAs and APs is mediated by ssDNA. The use of ethidium bromide in the histidine-tag pull-downs (Lai and Herr 1992) would help to determine if the RPA:AP interaction is DNA-independent, or whether the RPA:AP complex is formed upon binding ssDNA. Ethidium bromide, a DNA intercalator, distorts the DNA helix and consequently disrupts any protein:DNA interactions in a non-specific manner. Therefore addition of ethidium bromide to the cell lysate is considered a straightforward and highly efficient way to determine if protein:protein interactions are DNA-dependent or DNA independent (Lai and Herr 1992). The next important question regarding the RPA:AP interaction concerns the structure of these complexes, which would be resolved by obtaining the crystal structure of the RPA:AP complexes. Due to the halophilic nature of RPA1 and 3 and their respective associated proteins, the salt conditions in any crystallization trials would need to be considered carefully (Karen Bunting personal communication). However, the N-terminal histidine-tag may be affecting the structure of RPA3AP, resulting in a breakdown product of RPA3AP and decreased RPA3AP:RPA3 interactions. The construction of a C-terminal histidine-tagged RPA3AP might overcome this hindrance.

The mild DNA damage sensitivity of the double  $\Delta rpa1$  operon  $\Delta rpa3$  operon mutant, compared to that of the single  $\Delta rpa3$  operon mutant, and the absence of a growth defect was unexpected when considering that very few potential pop-outs were generated and that the *in trans* expression of the *rpa1* operon was required. The presence of a spontaneous suppressor mutation in the double  $\Delta rpa1$  operon  $\Delta rpa3$  operon mutant might explain the mild phenotype, even though the mutant was challenging to make. Therefore the examination of the double *rpa1 rpa3* operon deletion mutant for a suppressor mutation is necessary to establish if the *rpa1* operon and the *rpa3* operon share a function and therefore are redundant with each other. The sequencing of the *rpa2* gene would address whether mutations in RPA2 are compensating for the absence of the *rpa1* and *rpa3* operon, for example as a result of

increased affinity for ssDNA or merely due to overexpression of RPA2. Such mutations have been observed in *H. salinarum*, where elevated levels of expression of RPA can lead to increased resistance to ionizing radiation (DeVeaux, Muller *et al.* 2007). Furthermore, the attempt to replace *rpa2* with a copy of *rpa2* under an inducible promoter, using a new concept that includes positive selection at the pop-in and pop-out stage, will hopefully answer the question of whether RPA2 functions in DNA replication or repair or both.

There is a unifying theme in archaea of great variety in the number and type of proteins involved in DNA replication, and repair, whose counterparts in eukaryotes are much more uniform. This has been shown to be the case for RPA, where eukaryotes possess three subunits that form unified clades in a phylogenetic analysis, but in archaea the number and structure of subunits varies widely. Euryarchaea possess differing numbers of RPA subunits and some possess differing numbers of RPAs and APs. Crenarchaea also possess varying numbers of single-stranded DNA-binding proteins (SSB), however in both euryarchaea and crenarchaea none of the RPAs, APs or SSBs fall into unified clades. Again this is seen in the case for MCM, where eukaryotes possess six MCM subunits that each form unified clades. In archaea there is a vast range in the number of MCM subunits, which differ between individual species, but none of which fall into uniform clades. Characterising the RPA-AP complexes of *H. volcanii* will shed light on how the RPAs and APs function together in binding and stabilising ssDNA providing protection from degradation. This in turn will provide insight for other RPAs and APs in archaea, but also offer reasoning behind the driving force of such a non-uniform evolution of archaea.

## Bibliography

- Akita, M., A. Adachi, *et al.* (2010). Cdc6/Orc1 from *Pyrococcus furiosus* may act as the origin recognition protein and Mcm helicase recruiter. *Genes to cells : devoted to molecular & cellular mechanisms* **15**(5): 537-552.
- Alberts, B. (2003). DNA replication and recombination. *Nature* **421**(6921): 431-435.
- Alberts, B. M. and L. Frey (1970). T4 bacteriophage gene 32: a structural protein in the replication and recombination of DNA. *Nature* **227**(5265): 1313-1318.
- Allers, T., S. Barak, *et al.* (2010). Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in *Haloferax volcanii*. *Applied and environmental microbiology* **76**(6): 1759-1769.
- Allers, T. and M. Mevarech (2005). Archaeal genetics - the third way. *Nat Rev Genet* **6**(1): 58-73.
- Allers, T. and H. P. Ngo (2003). Genetic analysis of homologous recombination in Archaea: *Haloferax volcanii* as a model organism. *Biochem Soc Trans* **31**(Pt 3): 706-710.
- Allers, T., H. P. Ngo, *et al.* (2004). Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl Environ Microbiol* **70**(2): 943-953.
- Anantha, R. W., E. Sokolova, *et al.* (2008). RPA phosphorylation facilitates mitotic exit in response to mitotic DNA damage. *Proc Natl Acad Sci U S A* **105**(35): 12903-12908.
- Aparicio, O. M., D. M. Weinstein, *et al.* (1997). Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. *Cell* **91**(1): 59-69.
- Bailey, S., W. K. Eliason, *et al.* (2007). Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science* **318**(5849): 459-463.
- Barry, E. R. and S. D. Bell (2006). DNA replication in the archaea. *Microbiol Mol Biol Rev* **70**(4): 876-887.
- Bell, C. E. (2005). Structure and mechanism of *Escherichia coli* RecA ATPase. *Molecular microbiology* **58**(2): 358-366.
- Bell, S. D. (2011). DNA replication: archaeal oriGINS. *BMC biology* **9**: 36.
- Bell, S. P. (1995). Eukaryotic replicators and associated protein complexes. *Current opinion in genetics & development* **5**(2): 162-167.
- Berthon, J., D. Cortez, *et al.* (2008). Genomic context analysis in Archaea suggests previously unrecognized links between DNA replication and translation. *Genome biology* **9**(4): R71.
- Biswas, S. and D. Bastia (2008). Mechanistic insights into replication termination as revealed by investigations of the Reb1-Ter3 complex of *Schizosaccharomyces pombe*. *Molecular and cellular biology* **28**(22): 6844-6857.
- Bitan-Banin, G., R. Ortenberg, *et al.* (2003). Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the *pyrE* gene. *J Bacteriol* **185**(3): 772-778.
- Blanco, M. G., J. Matos, *et al.* (2010). Functional overlap between the structure-specific nucleases Yen1 and Mus81-Mms4 for DNA-damage repair in *S. cerevisiae*. *DNA repair* **9**(4): 394-402.
- Bochkarev, A. and E. Bochkareva (2004). From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr Opin Struct Biol* **14**(1): 36-42.
- Bochkareva, E., S. Korolev, *et al.* (2000). The role for zinc in replication protein A. *J Biol Chem* **275**(35): 27332-27338.

- Bochkareva, E., S. Korolev, *et al.* (2002). Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J* **21**(7): 1855-1863.
- Boddy, M. N., P. H. Gaillard, *et al.* (2001). Mus81-Eme1 are essential components of a Holliday junction resolvase. *Cell* **107**(4): 537-548.
- Boiteux, S. and M. Guillet (2004). Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. *DNA repair* **3**(1): 1-12.
- Bramhill, D. and A. Kornberg (1988). A model for initiation at origins of DNA replication. *Cell* **54**(7): 915-918.
- Braun, K. A., Y. Lao, *et al.* (1997). Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry* **36**(28): 8443-8454.
- Breuert, S., T. Allers, *et al.* (2006). Regulated polyploidy in halophilic archaea. *PLoS ONE* **1**: e92.
- Brewer, B. J. and W. L. Fangman (1988). A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell* **55**(4): 637-643.
- Brill, S. J. and B. Stillman (1991). Replication factor-A from *Saccharomyces cerevisiae* is encoded by three essential genes coordinately expressed at S phase. *Genes Dev* **5**(9): 1589-1600.
- Brochier-Armanet, C., P. Forterre, *et al.* (2011). Phylogeny and evolution of the Archaea: one hundred genomes later. *Current opinion in microbiology* **14**(3): 274-281.
- Bugreev, D. V., F. Hanaoka, *et al.* (2007). Rad54 dissociates homologous recombination intermediates by branch migration. *Nature structural & molecular biology* **14**(8): 746-753.
- Bujalowski, W., L. B. Overman, *et al.* (1988). Binding mode transitions of *Escherichia coli* single strand binding protein-single-stranded DNA complexes. Cation, anion, pH, and binding density effects. *J Biol Chem* **263**(10): 4629-4640.
- Bullock, W. O., J. M. Fernandez, *et al.* (1987). XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactocidase selection. *Biotechniques* **5**(4): 376-379.
- Cai, L., M. Roginskaya, *et al.* (2007). Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler. *Biochemistry* **46**(28): 8226-8233.
- Cann, I. K. and Y. Ishino (1999). Archaeal DNA replication. Identifying the pieces to solve a puzzle. *Genetics* **152**(4): 1249-1267.
- Castrec, B., C. Rouillon, *et al.* (2009). Binding to PCNA in Euryarchaeal DNA Replication requires two PIP motifs for DNA polymerase D and one PIP motif for DNA polymerase B. *Journal of molecular biology* **394**(2): 209-218.
- Charlebois, R. L., L. C. Schalkwyk, *et al.* (1991). Detailed physical map and set of overlapping clones covering the genome of the archaeobacterium *Haloferax volcanii* DS2. *J Mol Biol* **222**(3): 509-524.
- Chase, J. W., J. J. L'Italien, *et al.* (1984). Characterization of the *Escherichia coli* SSB-113 mutant single-stranded DNA-binding protein. Cloning of the gene, DNA and protein sequence analysis, high pressure liquid chromatography peptide mapping, and DNA-binding studies. *J Biol Chem* **259**(2): 805-814.
- Chia, N., I. Cann, *et al.* (2010). Evolution of DNA replication protein complexes in eukaryotes and Archaea. *PloS one* **5**(6): e10866.
- Christian, J. H. and J. A. Waltho (1962). Solute concentrations within cells of halophilic and non-halophilic bacteria. *Biochimica et biophysica acta* **65**: 506-508.



- Chuang, R. Y. and T. J. Kelly (1999). The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proceedings of the National Academy of Sciences of the United States of America* **96**(6): 2656-2661.
- Clark, A. B., F. Valle, *et al.* (2000). Functional interaction of proliferating cell nuclear antigen with MSH2-MSH6 and MSH2-MSH3 complexes. *The Journal of biological chemistry* **275**(47): 36498-36501.
- Cline, S. W., W. L. Lam, *et al.* (1989). Transformation methods for halophilic archaeobacteria. *Can J Microbiol* **35**(1): 148-152.
- Cline, S. W., L. C. Schalkwyk, *et al.* (1989). Transformation of the archaeobacterium *Halobacterium volcanii* with genomic DNA. *J Bacteriol* **171**(9): 4987-4991.
- Codlin, S. and J. Z. Dalgaard (2003). Complex mechanism of site-specific DNA replication termination in fission yeast. *The EMBO journal* **22**(13): 3431-3440.
- Coker, J. A., P. DasSarma, *et al.* (2009). Multiple replication origins of *Halobacterium* sp. strain NRC-1: properties of the conserved *orc7*-dependent *oriC1*. *Journal of bacteriology* **191**(16): 5253-5261.
- Coverley, D., M. K. Kenny, *et al.* (1992). A role for the human single-stranded DNA binding protein HSSB/RPA in an early stage of nucleotide excision repair. *Nucleic Acids Res* **20**(15): 3873-3880.
- Coverley, D., M. K. Kenny, *et al.* (1991). Requirement for the replication protein SSB in human DNA excision repair. *Nature* **349**(6309): 538-541.
- Crowley, D. J., I. Boubriak, *et al.* (2006). The *uvrA*, *uvrB* and *uvrC* genes are required for repair of ultraviolet light induced DNA photoproducts in *Halobacterium* sp. NRC-1. *Saline Systems* **2**: 11.
- Cubeddu, L. and M. F. White (2005). DNA damage detection by an archaeal single-stranded DNA-binding protein. *J Mol Biol* **353**(3): 507-516.
- Cullmann, G., K. Fien, *et al.* (1995). Characterization of the five replication factor C genes of *Saccharomyces cerevisiae*. *Molecular and cellular biology* **15**(9): 4661-4671.
- Cunningham, R. P., S. M. Saporito, *et al.* (1986). Endonuclease IV (*nfo*) mutant of *Escherichia coli*. *Journal of bacteriology* **168**(3): 1120-1127.
- Curth, U., J. Genschel, *et al.* (1996). In vitro and in vivo function of the C-terminus of *Escherichia coli* single-stranded DNA binding protein. *Nucleic Acids Res* **24**(14): 2706-2711.
- Daiyasu, H., K. Komori, *et al.* (2000). Hjc resolvase is a distantly related member of the type II restriction endonuclease family. *Nucleic Acids Res* **28**(22): 4540-4543.
- Daley, J. M., P. L. Palmbo, *et al.* (2005). Nonhomologous end joining in yeast. *Annual review of genetics* **39**: 431-451.
- Davey, M. J., L. Fang, *et al.* (2002). The DnaC helicase loader is a dual ATP/ADP switch protein. *EMBO J* **21**(12): 3148-3159.
- David, S. S. (2005). Structural biology: DNA search and rescue. *Nature* **434**(7033): 569-570.
- Delgado, S., M. Gomez, *et al.* (1998). Initiation of DNA replication at CpG islands in mammalian chromosomes. *The EMBO journal* **17**(8): 2426-2435.
- Delmas, S., L. Shunburne, *et al.* (2009). Mre11-Rad50 promotes rapid repair of DNA damage in the polyploid archaeon *Haloferax volcanii* by restraining homologous recombination. *PLoS genetics* **5**(7): e1000552.
- DeLong, E. F. and N. R. Pace (2001). Environmental diversity of bacteria and archaea. *Syst Biol* **50**(4): 470-478.

- Deng, X., J. E. Habel, *et al.* (2007). Structure of the full-length human RPA14/32 complex gives insights into the mechanism of DNA binding and complex formation. *J Mol Biol* **374**(4): 865-876.
- DePamphilis, M. L., J. J. Blow, *et al.* (2006). Regulating the licensing of DNA replication origins in metazoa. *Curr Opin Cell Biol* **18**(3): 231-239.
- DeVeaux, L. C., J. A. Muller, *et al.* (2007). Extremely radiation-resistant mutants of a halophilic archaeon with increased single-stranded DNA-binding protein (RPA) gene expression. *Radiation research* **168**(4): 507-514.
- Diffley, J. F. and K. Labib (2002). The chromosome replication cycle. *Journal of cell science* **115**(Pt 5): 869-872.
- Dillingham, M. S. and S. C. Kowalczykowski (2008). RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol Mol Biol Rev* **72**(4): 642-671, Table of Contents.
- Din, S., S. J. Brill, *et al.* (1990). Cell-cycle-regulated phosphorylation of DNA replication factor A from human and yeast cells. *Genes Dev* **4**(6): 968-977.
- Dionne, I., R. K. Nookala, *et al.* (2003). A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Molecular cell* **11**(1): 275-282.
- Donaldson, J. R., C. T. Courcelle, *et al.* (2006). RuvABC is required to resolve holliday junctions that accumulate following replication on damaged templates in *Escherichia coli*. *The Journal of biological chemistry* **281**(39): 28811-28821.
- Downs, J. A. and S. P. Jackson (2004). A means to a DNA end: the many roles of Ku. *Nature reviews. Molecular cell biology* **5**(5): 367-378.
- Duan, Z. (2008). *Genetic Analysis of Structure-Specific Endonucleases HEF and FEN1 in Archaeon Haloferax volcanii*
- Duderstadt, K. E., K. Chuang, *et al.* (2011). DNA stretching by bacterial initiators promotes replication origin opening. *Nature* **478**(7368): 209-213.
- Duggin, I. G., N. Dubarry, *et al.* (2011). Replication termination and chromosome dimer resolution in the archaeon *Sulfolobus solfataricus*. *The EMBO journal* **30**(1): 145-153.
- Duggin, I. G., S. A. McCallum, *et al.* (2008). Chromosome replication dynamics in the archaeon *Sulfolobus acidocaldarius*. *Proceedings of the National Academy of Sciences of the United States of America* **105**(43): 16737-16742.
- Duggin, I. G., R. G. Wake, *et al.* (2008). The replication fork trap and termination of chromosome replication. *Molecular microbiology* **70**(6): 1323-1333.
- Dutta, A. and B. Stillman (1992). cdc2 family kinases phosphorylate a human cell DNA replication factor, RPA, and activate DNA replication. *EMBO J* **11**(6): 2189-2199.
- El Yacoubi, B., G. Phillips, *et al.* (2009). A Gateway platform for functional genomics in *Haloferax volcanii*: deletion of three tRNA modification genes. *Archaea* **2**(4): 211-219.
- Erdile, L. F., M. S. Wold, *et al.* (1990). The primary structure of the 32-kDa subunit of human replication protein A. *J Biol Chem* **265**(6): 3177-3182.
- Evrin, C., P. Clarke, *et al.* (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proceedings of the National Academy of Sciences of the United States of America* **106**(48): 20240-20245.
- Fang, L., M. J. Davey, *et al.* (1999). Replisome assembly at oriC, the replication origin of *E. coli*, reveals an explanation for initiation sites outside an origin. *Mol Cell* **4**(4): 541-553.

- Fanning, E., V. Klimovich, *et al.* (2006). A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res* **34**(15): 4126-4137.
- Forget, A. L. and S. C. Kowalczykowski (2010). Single-molecule imaging brings Rad51 nucleoprotein filaments into focus. *Trends in cell biology* **20**(5): 269-276.
- Forsburg, S. L. (2004). Eukaryotic MCM proteins: beyond replication initiation. *Microbiology and molecular biology reviews : MMBR* **68**(1): 109-131.
- Fousteri, M. and L. H. Mullenders (2008). Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. *Cell research* **18**(1): 73-84.
- Freidberg, E. C., G. C. Walker, *et al.* (2006). *DNA Repair and Mutagenesis*.
- Frick, D. N. and C. C. Richardson (2001). DNA primases. *Annu Rev Biochem* **70**: 39-80.
- Friedman, J. I. and J. T. Stivers (2010). Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry* **49**(24): 4957-4967.
- Fuller, R. S., B. E. Funnell, *et al.* (1984). The dnaA protein complex with the E. coli chromosomal replication origin (oriC) and other DNA sites. *Cell* **38**(3): 889-900.
- Gambus, A., R. C. Jones, *et al.* (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol* **8**(4): 358-366.
- Garg, P., C. M. Stith, *et al.* (2004). Idling by DNA polymerase delta maintains a ligatable nick during lagging-strand DNA replication. *Genes & development* **18**(22): 2764-2773.
- Gari, K., C. Decaillet, *et al.* (2008). The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks. *Molecular cell* **29**(1): 141-148.
- Gaudier, M., B. S. Schuwirth, *et al.* (2007). Structural basis of DNA replication origin recognition by an ORC protein. *Science* **317**(5842): 1213-1216.
- Gellert, M., M. H. O'Dea, *et al.* (1976). Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. *Proceedings of the National Academy of Sciences of the United States of America* **73**(12): 4474-4478.
- Gilbert, D. M. (2001). Making sense of eukaryotic DNA replication origins. *Science* **294**(5540): 96-100.
- Grabowski, B. and Z. Kelman (2003). Archaeal DNA replication: eukaryal proteins in a bacterial context. *Annu Rev Microbiol* **57**: 487-516.
- Grainger, W. H., C. Machon, *et al.* (2010). DnaB proteolysis in vivo regulates oligomerization and its localization at oriC in Bacillus subtilis. *Nucleic acids research* **38**(9): 2851-2864.
- Grove, D. E., S. Willcox, *et al.* (2005). Differential single-stranded DNA binding properties of the paralogous SsbA and SsbB proteins from Streptococcus pneumoniae. *J Biol Chem* **280**(12): 11067-11073.
- Gu, L., Y. Hong, *et al.* (1998). ATP-dependent interaction of human mismatch repair proteins and dual role of PCNA in mismatch repair. *Nucleic acids research* **26**(5): 1173-1178.
- Guo, C., T. S. Tang, *et al.* (2010). SnapShot: nucleotide excision repair. *Cell* **140**(5): 754-754 e751.
- Guo, S., S. R. Presnell, *et al.* (2004). Differential requirement for proliferating cell nuclear antigen in 5' and 3' nick-directed excision in human mismatch repair. *The Journal of biological chemistry* **279**(17): 16912-16917.

- Guy, C. P., S. Haldenby, *et al.* (2006). Interactions of RadB, a DNA repair protein in archaea, with DNA and ATP. *J Mol Biol* **358**(1): 46-56.
- Haldenby, S. (2007). *Genetic analysis of RadB, a paralogue of the archaeal Rad51/RecA homologue, RadA.*
- Haldenby, S., M. F. White, *et al.* (2009). RecA family proteins in archaea: RadA and its cousins. *Biochem Soc Trans* **37**(Pt 1): 102-107.
- Haring, S. J., T. D. Humphreys, *et al.* (2010). A naturally occurring human RPA subunit homolog does not support DNA replication or cell-cycle progression. *Nucleic Acids Res* **38**(3): 846-858.
- Hartman, A. L., C. Norais, *et al.* The genome of *Haloferax volcanii*: comparative, historical and novel perspectives on a model archaeon. *In preparation.*
- Hartman, A. L., C. Norais, *et al.* (2010). The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PloS one* **5**(3): e9605.
- Haseltine, C. A. and S. C. Kowalczykowski (2002). A distinctive single-strand DNA-binding protein from the Archaeon *Sulfolobus solfataricus*. *Mol Microbiol* **43**(6): 1505-1515.
- Hassold, T., H. Hall, *et al.* (2007). The origin of human aneuploidy: where we have been, where we are going. *Human molecular genetics* **16 Spec No. 2**: R203-208.
- Hawkins, M. (2009). *DNA replication origins in Haloferax volcanii.*
- Hawkins, M., R. Retkute, *et al.* (Submitted). Deep sequencing uncovers origin-independent DNA replication origins in *Haloferax volcanii*.
- Hayashi, I., K. Morikawa, *et al.* (1999). Specific interaction between DNA polymerase II (PolD) and RadB, a Rad51/Dmc1 homolog, in *Pyrococcus furiosus*. *Nucleic Acids Res* **27**(24): 4695-4702.
- He, Z., L. A. Henricksen, *et al.* (1995). RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. *Nature* **374**(6522): 566-569.
- Heichinger, C., C. J. Penkett, *et al.* (2006). Genome-wide characterization of fission yeast DNA replication origins. *The EMBO journal* **25**(21): 5171-5179.
- Henneke, G., J. P. Raffin, *et al.* (2000). The PCNA from *Thermococcus fumicolans* functionally interacts with DNA polymerase delta. *Biochem Biophys Res Commun* **276**(2): 600-606.
- Heyer, W. D., M. R. Rao, *et al.* (1990). An essential *Saccharomyces cerevisiae* single-stranded DNA binding protein is homologous to the large subunit of human RP-A. *EMBO J* **9**(7): 2321-2329.
- Hingorani, M. M. and M. O'Donnell (1998). ATP binding to the *Escherichia coli* clamp loader powers opening of the ring-shaped clamp of DNA polymerase III holoenzyme. *J Biol Chem* **273**(38): 24550-24563.
- Hlavin, E. M., M. B. Smeaton, *et al.* (2010). Initiation of DNA interstrand cross-link repair in mammalian cells. *Environmental and molecular mutagenesis* **51**(6): 604-624.
- Holliday, R. (1964). A mechanism for gene conversion in fungi. *Genet. Res.* **5**: 282-304.
- Holmes, M., F. Pfeifer, *et al.* (1994). Improved shuttle vectors for *Haloferax volcanii* including a dual- resistance plasmid. *Gene* **146**(1): 117-121.
- Holmes, M. L. and M. L. Dyall-Smith (1990). A plasmid vector with a selectable marker for halophilic archaeobacteria. *J Bacteriol* **172**(2): 756-761.
- Holmes, M. L. and M. L. Dyall-Smith (1991). Mutations in DNA gyrase result in novobiocin resistance in halophilic archaeobacteria. *J Bacteriol* **173**(2): 642-648.
- Holmes, M. L. and M. L. Dyall-Smith (2000). Sequence and expression of a halobacterial beta-galactosidase gene. *Mol Microbiol* **36**(1): 114-122.

- Holthausen, J. T., C. Wyman, *et al.* (2010). Regulation of DNA strand exchange in homologous recombination. *DNA repair* **9**(12): 1264-1272.
- Huet, J., R. Schnabel, *et al.* (1983). Archaeobacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. *EMBO J* **2**(8): 1291-1294.
- Indiani, C. and M. O'Donnell (2006). The replication clamp-loading machine at work in the three domains of life. *Nature reviews. Molecular cell biology* **7**(10): 751-761.
- Jacob, F. and S. Brenner (1963). [On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon]. *Comptes rendus hebdomadaires des seances de l'Academie des sciences* **256**: 298-300.
- Jacobs, D. M., A. S. Lipton, *et al.* (1999). Human replication protein A: global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker. *Journal of biomolecular NMR* **14**(4): 321-331.
- Jiricny, J. (2006). The multifaceted mismatch-repair system. *Nature reviews. Molecular cell biology* **7**(5): 335-346.
- Johnson, R. E., C. A. Torres-Ramos, *et al.* (1998). Identification of APN2, the *Saccharomyces cerevisiae* homolog of the major human AP endonuclease HAP1, and its role in the repair of abasic sites. *Genes & development* **12**(19): 3137-3143.
- Jung, D. and F. W. Alt (2004). Unraveling V(D)J recombination; insights into gene regulation. *Cell* **116**(2): 299-311.
- Kantake, N., T. Sugiyama, *et al.* (2003). The recombination-deficient mutant RPA (rfa1-t11) is displaced slowly from single-stranded DNA by Rad51 protein. *The Journal of biological chemistry* **278**(26): 23410-23417.
- Kasiviswanathan, R., J. H. Shin, *et al.* (2005). Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties. *Nucleic acids research* **33**(15): 4940-4950.
- Katayama, T. and K. Sekimizu (1999). Inactivation of *Escherichia coli* DnaA protein by DNA polymerase III and negative regulations for initiation of chromosomal replication. *Biochimie* **81**(8-9): 835-840.
- Kazmirski, S. L., Y. Zhao, *et al.* (2005). Out-of-plane motions in open sliding clamps: molecular dynamics simulations of eukaryotic and archaeal proliferating cell nuclear antigen. *Proceedings of the National Academy of Sciences of the United States of America* **102**(39): 13801-13806.
- Kee, Y. and A. D. D'Andrea (2010). Expanded roles of the Fanconi anemia pathway in preserving genomic stability. *Genes & development* **24**(16): 1680-1694.
- Kelman, Z. and M. O'Donnell (1995). DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. *Annu Rev Biochem* **64**: 171-200.
- Kenny, M. K., S. H. Lee, *et al.* (1989). Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases alpha and delta. *Proc Natl Acad Sci U S A* **86**(24): 9757-9761.
- Kerr, I. D., R. I. Wadsworth, *et al.* (2003). Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of *Sulfolobus* SSB protein. *EMBO J* **22**(11): 2561-2570.
- Keshav, K. F., C. Chen, *et al.* (1995). Rpa4, a homolog of the 34-kilodalton subunit of the replication protein A complex. *Mol Cell Biol* **15**(6): 3119-3128.
- Killelea, T., S. Ghosh, *et al.* (2010). Probing the interaction of archaeal DNA polymerases with deaminated bases using X-ray crystallography and non-hydrogen bonding isosteric base analogues. *Biochemistry* **49**(27): 5772-5781.

- Kim, C., R. O. Snyder, *et al.* (1992). Binding properties of replication protein A from human and yeast cells. *Mol Cell Biol* **12**(7): 3050-3059.
- Komori, K. and Y. Ishino (2001). Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination. *J Biol Chem* **276**(28): 25654-25660.
- Komori, K., T. Miyata, *et al.* (2000). Both RadA and RadB are involved in homologous recombination in *Pyrococcus furiosus*. *J Biol Chem* **275**(43): 33782-33790.
- Kornberg, A. and T. A. Baker (1992). *DNA Replication*. New York, W.H. Freeman and Co.
- Kozlov, A. G., M. M. Cox, *et al.* (2010). Regulation of single-stranded DNA binding by the C termini of *Escherichia coli* single-stranded DNA-binding (SSB) protein. *J Biol Chem* **285**(22): 17246-17252.
- Krokan, H. E., H. Nilsen, *et al.* (2000). Base excision repair of DNA in mammalian cells. *FEBS letters* **476**(1-2): 73-77.
- Kubota, Y., Y. Takase, *et al.* (2003). A novel ring-like complex of *Xenopus* proteins essential for the initiation of DNA replication. *Genes & development* **17**(9): 1141-1152.
- Kuriyan, J. and M. O'Donnell (1993). Sliding clamps of DNA polymerases. *J Mol Biol* **234**(4): 915-925.
- Lai, J. S. and W. Herr (1992). Ethidium bromide provides a simple tool for identifying genuine DNA-independent protein associations. *Proceedings of the National Academy of Sciences of the United States of America* **89**(15): 6958-6962.
- Lam, W. L. and W. F. Doolittle (1989). Shuttle vectors for the archaebacterium *Halobacterium volcanii*. *Proc Natl Acad Sci U S A* **86**(14): 5478-5482.
- Lam, W. L. and W. F. Doolittle (1992). Mevinolin-resistant mutations identify a promoter and the gene for a eukaryote-like 3-hydroxy-3-methylglutaryl-coenzyme A reductase in the archaebacterium *Haloferax volcanii*. *J Biol Chem* **267**(9): 5829-5834.
- Lanyi, J. K. (1974). Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriological reviews* **38**(3): 272-290.
- Lao-Sirieix, S. H., R. K. Nookala, *et al.* (2005). Structure of the heterodimeric core primase. *Nature structural & molecular biology* **12**(12): 1137-1144.
- Large, A., C. Stamme, *et al.* (2007). Characterization of a tightly controlled promoter of the halophilic archaeon *Haloferax volcanii* and its use in the analysis of the essential *cct1* gene. *Molecular microbiology* **66**(5): 1092-1106.
- Larson, M. A., M. A. Griep, *et al.* (2010). Class-specific restrictions define primase interactions with DNA template and replicative helicase. *Nucleic acids research* **38**(20): 7167-7178.
- Lehoczky, P., P. J. McHugh, *et al.* (2007). DNA interstrand cross-link repair in *Saccharomyces cerevisiae*. *FEMS microbiology reviews* **31**(2): 109-133.
- Leigh, J. A., S. V. Albers, *et al.* (2011). Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS microbiology reviews* **35**(4): 577-608.
- Lestini, R., Z. Duan, *et al.* (2010). The archaeal Xpf/Mus81/FANCM homolog Hef and the Holliday junction resolvase Hjc define alternative pathways that are essential for cell viability in *Haloferax volcanii*. *DNA repair* **9**(9): 994-1002.
- Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry* **79**: 181-211.

- Lin, Y. L., C. Chen, *et al.* (1996). Dissection of functional domains of the human DNA replication protein complex replication protein A. *J Biol Chem* **271**(29): 17190-17198.
- Lin, Y. L., M. K. Shivji, *et al.* (1998). The evolutionarily conserved zinc finger motif in the largest subunit of human replication protein A is required for DNA replication and mismatch repair but not for nucleotide excision repair. *J Biol Chem* **273**(3): 1453-1461.
- Lindahl, T. (2001). Keynote: past, present, and future aspects of base excision repair. *Progress in nucleic acid research and molecular biology* **68**: xvii-xxx.
- Lohman, T. M. and M. E. Ferrari (1994). Escherichia coli single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annu Rev Biochem* **63**: 527-570.
- Lohman, T. M. and L. B. Overman (1985). Two binding modes in Escherichia coli single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J Biol Chem* **260**(6): 3594-3603.
- Longhese, M. P., P. Plevani, *et al.* (1994). Replication factor A is required in vivo for DNA replication, repair, and recombination. *Mol Cell Biol* **14**(12): 7884-7890.
- Lu, D. and J. L. Keck (2008). Structural basis of Escherichia coli single-stranded DNA-binding protein stimulation of exonuclease I. *Proc Natl Acad Sci U S A* **105**(27): 9169-9174.
- Lu, D., M. A. Windsor, *et al.* (2009). Peptide inhibitors identify roles for SSB C-terminal residues in SSB/exonuclease I complex formation. *Biochemistry* **48**(29): 6764-6771.
- Lundgren, M., A. Andersson, *et al.* (2004). Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci U S A* **101**(18): 7046-7051.
- Macneill, S. A. (2009). The haloarchaeal chromosome replication machinery. *Biochem Soc Trans* **37**(Pt 1): 108-113.
- Maisnier-Patin, S., L. Malandrin, *et al.* (2002). Chromosome replication patterns in the hyperthermophilic euryarchaea *Archaeoglobus fulgidus* and *Methanocaldococcus* (*Methanococcus*) *jannaschii*. *Mol Microbiol* **45**(5): 1443-1450.
- Marg, B. L., K. Schweimer, *et al.* (2005). A two-alpha-helix extra domain mediates the halophilic character of a plant-type ferredoxin from halophilic archaea. *Biochemistry* **44**(1): 29-39.
- Marti, T. M., C. Kunz, *et al.* (2002). DNA mismatch repair and mutation avoidance pathways. *Journal of cellular physiology* **191**(1): 28-41.
- Matsunaga, T., C. H. Park, *et al.* (1996). Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease. *J Biol Chem* **271**(19): 11047-11050.
- McCabe, K. M., S. B. Olson, *et al.* (2009). DNA interstrand crosslink repair in mammalian cells. *Journal of cellular physiology* **220**(3): 569-573.
- McCready, S. and L. Marcello (2003). Repair of UV damage in *Halobacterium salinarum*. *Biochem Soc Trans* **31**(Pt 3): 694-698.
- Mechali, M. (2010). Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat Rev Mol Cell Biol* **11**(10): 728-738.
- Meetei, A. R., A. L. Medhurst, *et al.* (2005). A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet* **37**(9): 958-963.

- Messer, W. (2002). The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev* **26**(4): 355-374.
- Mevarech, M. and R. Werczberger (1985). Genetic transfer in *Halobacterium volcanii*. *J Bacteriol* **162**(1): 461-462.
- Meyer, R. R. and P. S. Laine (1990). The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol Rev* **54**(4): 342-380.
- Mimitou, E. P. and L. S. Symington (2011). DNA end resection--unraveling the tail. *DNA repair* **10**(3): 344-348.
- Mimura, S. and H. Takisawa (1998). Xenopus Cdc45-dependent loading of DNA polymerase alpha onto chromatin under the control of S-phase Cdk. *The EMBO journal* **17**(19): 5699-5707.
- Miyata, T., H. Suzuki, *et al.* (2005). Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. *Proceedings of the National Academy of Sciences of the United States of America* **102**(39): 13795-13800.
- Moldovan, G. L., B. Pfander, *et al.* (2007). PCNA, the maestro of the replication fork. *Cell* **129**(4): 665-679.
- Moore, S. P., L. Erdile, *et al.* (1991). The human homologous pairing protein HPP-1 is specifically stimulated by the cognate single-stranded binding protein hRP-A. *Proc Natl Acad Sci U S A* **88**(20): 9067-9071.
- Moreau, P. L. (1988). Overproduction of single-stranded-DNA-binding protein specifically inhibits recombination of UV-irradiated bacteriophage DNA in *Escherichia coli*. *J Bacteriol* **170**(6): 2493-2500.
- Mott, M. L., J. P. Erzberger, *et al.* (2008). Structural synergy and molecular crosstalk between bacterial helicase loaders and replication initiators. *Cell* **135**(4): 623-634.
- Moyer, S. E., P. W. Lewis, *et al.* (2006). Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci U S A* **103**(27): 10236-10241.
- Mulcair, M. D., P. M. Schaeffer, *et al.* (2006). A molecular mousetrap determines polarity of termination of DNA replication in *E. coli*. *Cell* **125**(7): 1309-1319.
- Mullakhanbhai, M. F. and H. Larsen (1975). *Halobacterium volcanii* spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch Microbiol* **104**(3): 207-214.
- Myllykallio, H., P. Lopez, *et al.* (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* **288**(5474): 2212-2215.
- Nagata, Y., K. Mashimo, *et al.* (2002). The roles of Klenow processing and flap processing activities of DNA polymerase I in chromosome instability in *Escherichia coli* K12 strains. *Genetics* **160**(1): 13-23.
- Nagelhus, T. A., T. Haug, *et al.* (1997). A sequence in the N-terminal region of human uracil-DNA glycosylase with homology to XPA interacts with the C-terminal part of the 34-kDa subunit of replication protein A. *J Biol Chem* **272**(10): 6561-6566.
- Newlon, C. S. and J. F. Theis (1993). The structure and function of yeast ARS elements. *Current opinion in genetics & development* **3**(5): 752-758.
- Niedernhofer, L. J., A. S. Lalai, *et al.* (2005). Fanconi anemia (cross)linked to DNA repair. *Cell* **123**(7): 1191-1198.
- Nishino, T., K. Komori, *et al.* (2001). Dissection of the regional roles of the archaeal Holliday junction resolvase Hjc by structural and mutational analyses. *J Biol Chem* **276**(38): 35735-35740.
- Norais, C., M. Hawkins, *et al.* (2007). Genetic and physical mapping of DNA replication origins in *Haloferax volcanii*. *PLoS Genet* **3**(5): e77.



- Okuno, Y., H. Satoh, *et al.* (1999). Clustered adenine/thymine stretches are essential for function of a fission yeast replication origin. *Molecular and cellular biology* **19**(10): 6699-6709.
- Onrust, R., J. Finkelstein, *et al.* (1995). Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. III. Interface between two polymerases and the clamp loader. *J Biol Chem* **270**(22): 13366-13377.
- Oren, A. (2008). Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline systems* **4**: 2.
- Otterlei, M., E. Warbrick, *et al.* (1999). Post-replicative base excision repair in replication foci. *EMBO J* **18**(13): 3834-3844.
- Outwin, E., G. Carpenter, *et al.* (2011). Increased RPA1 Gene Dosage Affects Genomic Stability Potentially Contributing to 17p13.3 Duplication Syndrome. *PLoS genetics* **7**(8): e1002247.
- Oyama, M., M. Wakasugi, *et al.* (2004). Human NTH1 physically interacts with p53 and proliferating cell nuclear antigen. *Biochemical and biophysical research communications* **321**(1): 183-191.
- Ozaki, S. and T. Katayama (2011). Highly organized DnaA-oriC complexes recruit the single-stranded DNA for replication initiation. *Nucleic acids research*.
- Pacek, M., A. V. Tutter, *et al.* (2006). Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Molecular cell* **21**(4): 581-587.
- Palomera-Sanchez, Z. and M. Zurita (2011). Open, repair and close again: chromatin dynamics and the response to UV-induced DNA damage. *DNA repair* **10**(2): 119-125.
- Pan, M., T. J. Santangelo, *et al.* (2011). Thermococcus kodakarensis encodes three MCM homologs but only one is essential. *Nucleic acids research*.
- Park, M. S., D. L. Ludwig, *et al.* (1996). Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J Biol Chem* **271**(31): 18996-19000.
- Peterson, S. N., C. K. Sung, *et al.* (2004). Identification of competence pheromone responsive genes in Streptococcus pneumoniae by use of DNA microarrays. *Mol Microbiol* **51**(4): 1051-1070.
- Pitcher, R. S., N. C. Brissett, *et al.* (2007). Nonhomologous end-joining in bacteria: a microbial perspective. *Annual review of microbiology* **61**: 259-282.
- Raghunathan, S., A. G. Kozlov, *et al.* (2000). Structure of the DNA binding domain of E. coli SSB bound to ssDNA. *Nature structural biology* **7**(8): 648-652.
- Remus, D., F. Beuron, *et al.* (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**(4): 719-730.
- Remus, D. and J. F. Diffley (2009). Eukaryotic DNA replication control: lock and load, then fire. *Current opinion in cell biology* **21**(6): 771-777.
- Reuter, C. J. and J. Maupin-Furlow (2004). Analysis of proteasome-dependent proteolysis in Haloferax volcanii cells, using short-lived green fluorescent proteins. *Appl Environ Microbiol* **70**(12): 7530-7538.
- Richard, D. J., S. D. Bell, *et al.* (2004). Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase. *Nucleic Acids Res* **32**(3): 1065-1074.
- Ricke, R. M. and A. K. Bielinsky (2004). Mcm10 regulates the stability and chromatin association of DNA polymerase-alpha. *Molecular cell* **16**(2): 173-185.

- Rivera, M. C., R. Jain, *et al.* (1998). Genomic evidence for two functionally distinct gene classes. *Proc Natl Acad Sci U S A* **95**(11): 6239-6244.
- Robbins, J. B., M. C. Murphy, *et al.* (2004). Functional analysis of multiple single-stranded DNA-binding proteins from *Methanosarcina acetivorans* and their effects on DNA synthesis by DNA polymerase BI. *J Biol Chem* **279**(8): 6315-6326.
- Robertson, A. B., A. Klungland, *et al.* (2009). DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cellular and molecular life sciences : CMLS* **66**(6): 981-993.
- Robinson, J. L., B. Pyzyna, *et al.* (2005). Growth kinetics of extremely halophilic archaea (family halobacteriaceae) as revealed by arrhenius plots. *Journal of bacteriology* **187**(3): 923-929.
- Robinson, N. P. and S. D. Bell (2005). Origins of DNA replication in the three domains of life. *FEBS J* **272**(15): 3757-3766.
- Rolfsmeier, M. L. and C. A. Haseltine (2010). The single-stranded DNA binding protein of *Sulfolobus solfataricus* acts in the presynaptic step of homologous recombination. *J Mol Biol* **397**(1): 31-45.
- Rosenshine, I. and M. Mevarech (1989). Isolation and partial characterization of plasmids found in three *Halobacterium volcanii* isolates. *Can. J. Microbiol.* **35**: 92-95.
- Rouillon, C. and M. F. White (2011). The evolution and mechanisms of nucleotide excision repair proteins. *Research in microbiology* **162**(1): 19-26.
- Rudolph, C. J., A. L. Upton, *et al.* (2010). Is RecG a general guardian of the bacterial genome? *DNA repair* **9**(3): 210-223.
- San Filippo, J., P. Sung, *et al.* (2008). Mechanism of eukaryotic homologous recombination. *Annual review of biochemistry* **77**: 229-257.
- Schaper, S. and W. Messer (1995). Interaction of the initiator protein DnaA of *Escherichia coli* with its DNA target. *J Biol Chem* **270**(29): 17622-17626.
- Schwartz, E. K. and W. D. Heyer (2011). Processing of joint molecule intermediates by structure-selective endonucleases during homologous recombination in eukaryotes. *Chromosoma* **120**(2): 109-127.
- Sclafani, R. A. and T. M. Holzen (2007). Cell cycle regulation of DNA replication. *Annual review of genetics* **41**: 237-280.
- Segurado, M., A. de Luis, *et al.* (2003). Genome-wide distribution of DNA replication origins at A+T-rich islands in *Schizosaccharomyces pombe*. *EMBO reports* **4**(11): 1048-1053.
- Seitz, E. M., J. P. Brockman, *et al.* (1998). RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev* **12**(9): 1248-1253.
- Seybert, A., M. R. Singleton, *et al.* (2006). Communication between subunits within an archaeal clamp-loader complex. *The EMBO journal* **25**(10): 2209-2218.
- Shereda, R. D., A. G. Kozlov, *et al.* (2008). SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* **43**(5): 289-318.
- Shereda, R. D., N. J. Reiter, *et al.* (2009). Identification of the SSB binding site on *E. coli* RecQ reveals a conserved surface for binding SSB's C terminus. *J Mol Biol* **386**(3): 612-625.
- Shin, D. S., L. Pellegrini, *et al.* (2003). Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *EMBO J* **22**(17): 4566-4576.
- Shin, J. H. and Z. Kelman (2006). The replicative helicases of bacteria, archaea, and eukarya can unwind RNA-DNA hybrid substrates. *The Journal of biological chemistry* **281**(37): 26914-26921.

- Shinohara, A., M. Shinohara, *et al.* (1998). Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. *Genes to cells : devoted to molecular & cellular mechanisms* **3**(3): 145-156.
- Sigal, N., H. Delius, *et al.* (1972). A DNA-unwinding protein isolated from *Escherichia coli*: its interaction with DNA and with DNA polymerases. *Proc Natl Acad Sci U S A* **69**(12): 3537-3541.
- Singleton, M. R., M. S. Dillingham, *et al.* (2004). Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* **432**(7014): 187-193.
- Smith, P., P. A. Nair, *et al.* (2011). Structures and activities of archaeal members of the LigD 3'-phosphoesterase DNA repair enzyme superfamily. *Nucleic acids research* **39**(8): 3310-3320.
- Soppa, J. and D. Oesterhelt (1989). Bacteriorhodopsin mutants of *Halobacterium* sp. GRB. I. The 5-bromo-2'-deoxyuridine selection as a method to isolate point mutants in halobacteria. *The Journal of biological chemistry* **264**(22): 13043-13048.
- Spang, A., R. Hatzenpichler, *et al.* (2010). Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends in microbiology* **18**(8): 331-340.
- Speck, C. and W. Messer (2001). Mechanism of origin unwinding: sequential binding of DnaA to double- and single-stranded DNA. *EMBO J* **20**(6): 1469-1476.
- Speck, C., C. Weigel, *et al.* (1999). ATP- and ADP-dnaA protein, a molecular switch in gene regulation. *EMBO J* **18**(21): 6169-6176.
- Stigger, E., R. Drissi, *et al.* (1998). Functional analysis of human replication protein A in nucleotide excision repair. *J Biol Chem* **273**(15): 9337-9343.
- Stinchcomb, D. T., K. Struhl, *et al.* (1979). Isolation and characterisation of a yeast chromosomal replicator. *Nature* **282**(5734): 39-43.
- Struhl, K., D. T. Stinchcomb, *et al.* (1979). High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America* **76**(3): 1035-1039.
- Stukenberg, P. T., J. Turner, *et al.* (1994). An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps. *Cell* **78**(5): 877-887.
- Symington, L. S. (2002). Role of RAD52 Epistasis Group Genes in Homologous Recombination and Double-Strand Break Repair. *Microbiol Mol Biol Rev* **66**(4): 630-670.
- Takayama, Y., Y. Kamimura, *et al.* (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes & development* **17**(9): 1153-1165.
- Tomasz, M., R. Lipman, *et al.* (1987). Isolation and structure of a covalent cross-link adduct between mitomycin C and DNA. *Science* **235**(4793): 1204-1208.
- Turner, J., M. M. Hingorani, *et al.* (1999). The internal workings of a DNA polymerase clamp-loading machine. *EMBO J* **18**(3): 771-783.
- Van Houten, B., D. L. Croteau, *et al.* (2005). 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. *Mutation research* **577**(1-2): 92-117.
- Wallace, R. B., J. Shaffer, *et al.* (1979). Hybridization of synthetic oligodeoxyribonucleotides to phi chi 174 DNA: the effect of single base pair mismatch. *Nucleic acids research* **6**(11): 3543-3557.
- Walters, A. D. and J. P. Chong (2010). An archaeal order with multiple minichromosome maintenance genes. *Microbiology* **156**(Pt 5): 1405-1414.

- Walther, A. P., X. V. Gomes, *et al.* (1999). Replication protein A interactions with DNA. 1. Functions of the DNA-binding and zinc-finger domains of the 70-kDa subunit. *Biochemistry* **38**(13): 3963-3973.
- Wanner, C. and J. Soppa (1999). Genetic identification of three ABC transporters as essential elements for nitrate respiration in *Haloferax volcanii*. *Genetics* **152**(4): 1417-1428.
- Wendoloski, D., C. Ferrer, *et al.* (2001). A new simvastatin (mevinolin)-resistance marker from *Haloarcula hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2. *Microbiology* **147**(Pt 4): 959-964.
- White, M. F. (2011). Homologous recombination in the archaea: the means justify the ends. *Biochemical Society transactions* **39**(1): 15-19.
- White, M. F. and S. D. Bell (2002). Holding it together: chromatin in the Archaea. *Trends Genet* **18**(12): 621-626.
- Wobbe, C. R., L. Weissbach, *et al.* (1987). Replication of simian virus 40 origin-containing DNA in vitro with purified proteins. *Proc Natl Acad Sci U S A* **84**(7): 1834-1838.
- Woese, C. R. and G. E. Fox (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* **74**(11): 5088-5090.
- Woese, C. R., O. Kandler, *et al.* (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **87**(12): 4576-4579.
- Wold, M. S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66**: 61-92.
- Wold, M. S., D. H. Weinberg, *et al.* (1989). Identification of cellular proteins required for simian virus 40 DNA replication. *J Biol Chem* **264**(5): 2801-2809.
- Wolfe, R. S. (1996). 1776:1996 alessandro volta's combustible air. *ASM news*(62): 529-534.
- Wright, D. B., D. D. Banks, *et al.* (2002). The effect of salts on the activity and stability of *Escherichia coli* and *Haloferax volcanii* dihydrofolate reductases. *Journal of molecular biology* **323**(2): 327-344.
- Xie, P. and J. R. Sayers (2011). A model for transition of 5'-nuclease domain of DNA polymerase I from inert to active modes. *PLoS One* **6**(1): e16213.
- Yang, W. (2000). Structure and function of mismatch repair proteins. *Mutation research* **460**(3-4): 245-256.
- Yang, W., M. S. Junop, *et al.* (2000). DNA mismatch repair: from structure to mechanism. *Cold Spring Harbor symposia on quantitative biology* **65**: 225-232.
- Yao, N. Y. and M. O'Donnell (2010). SnapShot: The replisome. *Cell* **141**(6): 1088, 1088 e1081.
- Yeeles, J. T. and M. S. Dillingham (2010). The processing of double-stranded DNA breaks for recombinational repair by helicase-nuclease complexes. *DNA repair* **9**(3): 276-285.
- Yuzhakov, A., J. Turner, *et al.* (1996). Replisome assembly reveals the basis for asymmetric function in leading and lagging strand replication. *Cell* **86**(6): 877-886.
- Zhao, A., F. C. Gray, *et al.* (2006). ATP- and NAD-dependent DNA ligases share an essential function in the halophilic archaeon *Haloferax volcanii*. *Mol Microbiol* **59**(3): 743-752.
- Zhu, W., C. Ukomadu, *et al.* (2007). Mcm10 and And-1/CTF4 recruit DNA polymerase alpha to chromatin for initiation of DNA replication. *Genes & development* **21**(18): 2288-2299.

# Appendix

## RPA1 and RPA1AP MP2111

### 1A\_MP2111 RPA2

gene="rpa2"

locus\_tag="HVO\_0591"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110914/FtocobcOE.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110914/FtocobcOE.dat)

1. [gi|292654691](#) Mass: 51007 Score: 727 Matches: 12(6) Sequences: 11(6)  
replication protein A [Haloferax volcanii DS2]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/>	<u>9</u>	585.2925	1168.5704	1168.5836	-	0	68	0.012	1	U R.DGSLELHVGS.R.G
				0.0132						
<input checked="" type="checkbox"/>	<u>10</u>	601.3044	1200.5942	1200.6099	-	0	59	0.1	1	U R.VSNLSVGDP.TGR.V
				0.0156						
<input checked="" type="checkbox"/>	<u>11</u>	613.2864	1224.5582	1224.5622	-	0	60	0.066	1	U K.DGYNGVEVSASK.V
				0.0040						
<input checked="" type="checkbox"/>	<u>12</u>	634.7735	1267.5324	1267.5429	-	1	38	11	1	U R.TFDRDDGTEGR.V
				0.0105						
	<u>15</u>	651.3492	1300.6838	1300.6623	-	0	8	1.3e+04	6	U R.VVNIEVADETGR.I
				0.0216						
<input checked="" type="checkbox"/>	<u>17</u>	728.3533	1454.6920	1454.7001	-	0	87	0.00013	1	U R.ISAESVEVHTGAQQ.-
				0.0081						
<input checked="" type="checkbox"/>	<u>18</u>	728.3576	1454.7006	1454.7001	-	0	(20)	7.2e+02	1	U R.ISAESVEVHTGAQQ.-
				0.0005						
<input checked="" type="checkbox"/>	<u>19</u>	741.3335	1480.6524	1480.6543	-	1	51	0.59	1	U R.TFDRDDGSEGQVR.N
				0.0018						
<input checked="" type="checkbox"/>	<u>22</u>	836.4402	1670.8658	1670.8727	-	0	82	0.00043	1	U R.VEDLALGLSDVNLEGK.I
				0.0068						
<input checked="" type="checkbox"/>	<u>25</u>	1138.0565	2274.0984	2274.1016	-	0	113	2.8e-07	1	U R.ADLAELEPGQSVEVVDGYVR.E
				0.0031						
<input checked="" type="checkbox"/>	<u>27</u>	806.0499	2415.1279	2415.1410	-	0	72	0.003	1	U K.VEQMGLADEETAAMLIAHEL.R.D + 2
				0.0131						Oxidation (M)
<input checked="" type="checkbox"/>	<u>29</u>	1212.0530	2422.0914	2422.0984	-	0	91	3.5e-05	1	U K.TVDTDADLGLGDEVTVSGTETDGR.I
				0.0069						

Match to: [gi|292654691](#) Score: 727

replication protein A [Haloferax volcanii DS2]

Found in search of 1A\_MP2411\_14SEPT11.pk1

Nominal mass (M<sub>r</sub>): 51007; Calculated pI value: 3.96

NCBI BLAST search of [gi|292654691](#) against nr

Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Haloferax volcanii DS2](#)

Links to retrieve other entries containing this sequence from NCBI Entrez:

[gi|291372072](#) from [Haloferax volcanii DS2](#)

Variable modifications: Carbamidomethyl (C),Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 34%

Matched peptides shown in **Bold Red**

1 MGVIREVYDD LDTDFEFEEF EAAVNDK**VEQ MGGLADEETA AMLIAHEL**RD  
51 EEVNGIADIE PGMEDVKFLA KVVSIQEVRT FERDGEDEG **RVVNIEVADE**  
101 **TGR**IRVSLWD EMAAGAKENL EVGTVLRIGG RPK**DGYNGVE VSASKVEEDL**  
151 DAEVDVQVLD SYR**VEDLALG LSDVNLEGKI** LDAGTVR**TFD RDDGTEGRVS**  
201 **NLSVGDP**TGR VRVTLWDERA **DLAELEPGQ SVEVVDGYVR** ERDGSLELHV  
251 **GSRGAVEPID** EDIEYVPETT DIGSLELGQT VDIAGGVIEA DGKR**TFDRDD**  
301 **GSEGQVR**NIR VKDGTGDIRV ALWGEKADAD VDLADYVVIT DAEIKEGWQE  
351 DLEASAGWRS SVAVMDEAPE GAAGTDAGGS APTPPSDEGL GAFSGDGSSD  
401 DTSAANGGSS SDASAAESTG EAVEFTGTVV QAGTPVILDD GTQTK**TVDTD**  
451 **ADLGLGDEVT VSGTETDGR**I **SAESVEVHTG AQQ**

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110914/FtocobYTe.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110914/FtocobYTe.dat)

Sequence Coverage: 32%

Matched peptides shown in **Bold Red**

1	MGVIREVYDD	LDTDVEFEF	EAAVNDKVEQ	MGGLADEETA	AMLIAHEL	RD
51	EEVNGIADIE	PGMEDVKFLA	KVVSIGEVRT	FERDGEDEDG	RVVNIEVADE	
101	TGRIRVSLWD	EMAAGAKENL	EVGTVLIRIGG	RPKDGYNVE	VSASKVEEDL	
151	DAEVDVQVLD	SYRVEDLALG	LSDVNLEGKI	LDAGTVR	TFD	RDDGTEGRVS
201	NLSVGDP	TGRV	VRVTLWDERA	DLAEELEPGQ	SVEVVDGYVR	ERDGSLELHV
251	GSRGAVEPID	EDIEYVPETT	DIGSLELGQT	VDIAGGVIEA	DGKR	TFDRDD
301	GSEGQVRNIR	VKDGTGDIRV	ALWGEKADAD	VDLADYVVIT	DAEIK	KEGWQE
351	DLEASAGWRS	SVAVMDEAPE	GAAGTDAGGS	APTPPSDEGL	GAFSGDGSSD	
401	DTSAANGSS	SDASAAESTG	EAVEFTGTVV	QAGTPVILDD	GTQTK	TVDTD
451	ADLGLGDEVT	VSGTETDGRI	SAESVEVHTG	AQQ		

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
28 - 49	806.0499	2415.1279	2415.1410	-0.0131	0	K.VEQMGGLADEETAAMLIAHEL.R.D 2
Oxidation (M) (Ions score 72)						
134 - 145	613.2864	1224.5582	1225.5462	-0.9880	0	K.DGYNGVEVSASK.V Deamidated (NQ)
(Ions score 61)						
164 - 179	836.4402	1670.8658	1670.8727	-0.0068	0	R.VEDLALGLSDVNLEGK.I (Ions score 82)
188 - 198	634.7735	1267.5324	1267.5429	-0.0105	1	R.TFDRDDGTEGR.V (Ions score 38)
199 - 210	601.3044	1200.5942	1200.6099	-0.0156	0	R.VSNLSVGDP
220 - 240	1138.0565	2274.0984	2274.1016	-0.0031	0	R.ADLAEELEPGQSVEVVDGYVR.E (Ions score 113)
243 - 253	585.2925	1168.5704	1168.5836	-0.0132	0	R.DGSLELHVGSR.G (Ions score 68)
295 - 307	741.3335	1480.6524	1480.6543	-0.0018	1	R.TFDRDDGSEGQVR.N (Ions score 51)
446 - 469	1212.0530	2422.0914	2422.0984	-0.0069	0	K.TVDTDADLGLGDEVTVSGTETDGR.I (Ions score 91)
470 - 483	728.3533	1454.6920	1454.7001	-0.0081	0	R.ISAESVEVHTGAQQ.- (Ions score 87)
470 - 483	728.3576	1454.7006	1454.7001	0.0005	0	R.ISAESVEVHTGAQQ.- (Ions score 20)

# De novo peptides

836.44 +2 VEDLALGLSDVNLEGK residues 170-185

728 +2 LSAESVEVHTGAQQ confirms C-terminal peptide seq

## Combining results of all 1A searches and the blank run 1 peptides (46%):-

1	MGVIREVYDD	LDTDVEFEF	EAAVNDKVEQ	MGGLADEETA	AMLIAHEL	RD
51	EEVNGIADIE	PGMEDVKFLA	KVVSIGEVRT	FERDGEDEDG	RVVNIEVADE	
101	TGRIRVSLWD	EMAAGAKENL	EVGTVLIRIGG	RPKDGYNVE	VSASKVEEDL	
151	DAEVDVQVLD	SYRVEDLALG	LSDVNLEGKI	LDAGTVR	TFD	RDDGTEGRVS
201	NLSVGDP	TGRV	VRVTLWDERA	DLAEELEPGQ	SVEVVDGYVR	ERDGSLELHV
251	GSRGAVEPID	EDIEYVPETT	DIGSLELGQT	VDIAGGVIEA	DGKR	TFDRDD
301	GSEGQVRNIR	VKDGTGDIRV	ALWGEKADAD	VDLADYVVIT	DAEIK	KEGWQE
351	DLEASAGWRS	SVAVMDEAPE	GAAGTDAGGS	APTPPSDEGL	GAFSGDGSSD	
401	DTSAANGSS	SDASAAESTG	EAVEFTGTVV	QAGTPVILDD	GTQTK	TVDTD
451	ADLGLGDEVT	VSGTETDGRI	SAESVEVHTG	AQQ		

## Blank post 1A \_1

Rpa2 still present in the first blank.

Note there are peptides here not found in the actual run.

Note that most of the peptides here are all quite large (this makes sense – eluted later rather than during the short run).

Note – no sign of the N-terminal peptide in this run either.

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110914/FtocobcOh.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110914/FtocobcOh.dat)

1. [gi|292654691](#) Mass: 51007 Score: 220 Matches: 6(1) Sequences: 5(1)  
replication protein A [Haloferax volcanii DS2]

☐ Check to include this hit in error tolerant search

	Query	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/>	<a href="#">19</a>	817.3662	1632.7178	1632.7168	0.0010	0	28	98	1	U	K.EGWQEDLEASAGWR.S
<input checked="" type="checkbox"/>	<a href="#">20</a>	1018.5037	2034.9928	2034.9997	0.0069	0	6	1.4e+04	1	U	K.ADADVDLADYVVITDAEIK.E

<input checked="" type="checkbox"/>	22	1138.05762274.10062274.1016	0.0009	0	132	3.7e-09	1	U	R.ADLAEELEPGQSVVVDGYVR.E
<input checked="" type="checkbox"/>	24	800.7174 2399.13042399.1461	0.0157	0	(20)	4.5e+02	1	U	K.VEQMGGLADEETAAMLIAHEL.R.D + Oxidation (M)
<input checked="" type="checkbox"/>	25	806.0474 2415.12042415.1410	0.0206	0	24	1.8e+02	1	U	K.VEQMGGLADEETAAMLIAHEL.R.D + 2 Oxidation (M)
<input checked="" type="checkbox"/>	26	1296.56622591.11782591.1075	0.0103	0	29	51	1	U	R.EVYDDLDTDFEFEFEAAVNDK.V

# 1B\_MP2411 RPA1AP

gene="rpap1"

locus\_tag="HVO\_1337"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110915/FtocrmTOh.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110915/FtocrmTOh.dat)

1. gi|292655491 Mass: 64829 Score: 671 Matches: 12 (5) Sequences: 8 (5)  
rpa-associated protein [Haloferax volcanii DS2]

☐ Check to include this hit in error tolerant search

	Query	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/>	4	517.2624	1032.51021032	4.989	0.0114	0	57	0.2	1	U	R.AIDHYGTTR.T
<input checked="" type="checkbox"/>	7	571.8179	1141.62121141	6.091	0.0121	0	74	0.0032	1	U	R.GVDSTLAAGVPR.A
	8	608.8260	1215.63741215	6.611	0.0237	0	16	2e+03	3	U	R.WIVSAAEATLR.R
<input checked="" type="checkbox"/>	9	623.3363	1244.65801244	6.513	0.0067	0	55	0.25	1	U	R.APNYVVPTGAR.L
<input checked="" type="checkbox"/>	12	735.4072	1468.79981468	7.885	0.0113	0	97	1.5e-05	1	U	R.EVAVQALELVAGDR.D
<input checked="" type="checkbox"/>	14	756.3878	1510.76101510	7.490	0.0121	0	(15)	2e+03	1	U	R.IAVFDEALSMPYR.G
<input checked="" type="checkbox"/>	15	764.3871	1526.75961526	7.439	0.0158	0	(50)	0.64	1	U	R.IAVFDEALSMPYR.G + Oxidation (M)
<input checked="" type="checkbox"/>	16	764.3907	1526.76681526	7.439	0.0230	0	86	0.00015	1	U	R.IAVFDEALSMPYR.G + Oxidation (M)
<input checked="" type="checkbox"/>	18	1133.99302265	97142265	9.549	0.0165	0	(26)	1.2e+02	1	U	R.VFAAEFDDASLSYSESDEER.A
<input checked="" type="checkbox"/>	19	1134.00022265	98582265	9.549	0.0309	0	153	2.4e-11	1	U	R.VFAAEFDDASLSYSESDEER.A
<input checked="" type="checkbox"/>	20	1000.14232997	40512997	3.840	0.0211	0	(13)	1.7e+03	1	U	R.TYEPDDADVYSSVRPESVNTVDAD VR.D
<input checked="" type="checkbox"/>	21	1000.14542997	41442997	3.840	0.0304	0	135	1.2e-09	1	U	R.TYEPDDADVYSSVRPESVNTVDAD VR.D

Match to: gi|292655491 Score: 671

rpa-associated protein [Haloferax volcanii DS2]

Found in search of 1B\_MP2411\_14SEPT11.pkl

Nominal mass (M<sub>r</sub>): 64829; Calculated pI value: 3.49

NCBI BLAST search of gi|292655491 against nr

Unformatted sequence string for pasting into other applications

Taxonomy: Haloferax volcanii DS2

Links to retrieve other entries containing this sequence from NCBI Entrez:

gi|291371037 from Haloferax volcanii DS2

Variable modifications: Carbamidomethyl (C), Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 18%

Matched peptides shown in **Bold Red**

```

1 MSASPVGTR EIAYRVFAAE FDDASLSYSE SDEERAPNYV VTPTGARLNR
51 TFVAGVLTEV EHVNDVLRG RIADPTGAFV TYAGQYQPEP MAYLDAATPP
101 AFVSLAGKAR TYEPDDADV YSSVRPESVN TVDADVDRW IVSAAEATLR
151 RIAVFDEALS MPYRGDDLTR ALEARGVDST LAAGVPR HYGTTRTYLE
201 ALREEVAVQAL ELVAGDRDQV DPLDVAPGDG GDAVLGPLE LDLEPAASVD
251 IEVGGEADAGE VEADAGEFEN ESGSEPAATL DSEPEFDDGA ESEAVEPESE
301 SEPFAEPDSE PESASDPEPV AAETTAAAES QPGASESEPA SEPSTTVDS
351 PEPEPAEPE SEPAEPDLT SEPEPEPVT SLDTESEPDF DDGALGDFED
401 DAPDAGATDA GSLDDSGSGT LGDFDDGFDD PDPEAGDSTD STDSASADAD
451 DSVDPDGMYE LDEDERAEIE SEFGTEFTSG ADVDPAGEAD IDVPDADDLT
501 EQLEDESAAA AESDPEPATA AAEPEPAVDS EPEPTDAAS DGDEDEADAD
551 PDAESAEDI DLESVVVDAM DDLDDGDGAT RDEVVAAVD EHGADPGAVE
601 DAIQEALLGG RCYEPQDGV L KAI

```

# 1C\_MP2411 RPA1

SEARCHED FOR his tagged N-TERMINAL PEPTIDE, no sign of it  
 gene="rpa1"  
 locus\_tag="HVO\_1338"  
 ncbi

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110915/FtocrmTOe.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110915/FtocrmTOe.dat)

1. [gi|292655492](#) Mass: 45954 Score: 940 Matches: 15(10) Sequences: 11(8)  
 replication protein A [Haloferax volcanii DS2]

☐ Check to include this hit in error tolerant search

	Query	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/>	<u>1</u>	472.2861	942.5576	942.5498	0.0078	0	83	0.00051	1	U	R.AAALLTEVR.A
	<u>4</u>	693.7765	1385.5384	1386.5470	1.0086	0	(26)	1.7e+02	5	U	R.QHGEVDGEDDMR.V
<input checked="" type="checkbox"/>	<u>5</u>	702.2833	1402.5520	1402.5419	0.0101	0	60	0.076	1	U	R.QHGEVDGEDDMR.V + Oxidation (M)
<input checked="" type="checkbox"/>	<u>6</u>	704.8622	1407.7098	1407.6994	0.0104	0	74	0.0032	1	U	R.SIVYQGDEQTIR.E
<input checked="" type="checkbox"/>	<u>8</u>	736.8896	1471.7646	1471.7518	0.0128	0	60	0.065	1	U	K.EVVADDIADTLVGR.E
<input checked="" type="checkbox"/>	<u>9</u>	771.8969	1541.7792	1541.7685	0.0107	0	(76)	0.0016	1	U	R.IGGEADLIDLQAGDR.G
<input checked="" type="checkbox"/>	<u>10</u>	771.8980	1541.7814	1541.7685	0.0129	0	127	1.2e-08	1	U	R.IGGEADLIDLQAGDR.G
<input checked="" type="checkbox"/>	<u>11</u>	815.9458	1629.8770	1629.8574	0.0197	0	109	8.8e-07	1	U	K.AILDDGTGTVTAILDR.D
<input checked="" type="checkbox"/>	<u>12</u>	828.4410	1654.8674	1654.8526	0.0148	0	104	2.9e-06	1	U	R.IVDIDPDGGNVSVTVR.V
<input checked="" type="checkbox"/>	<u>13</u>	828.4435	1654.8724	1654.8526	0.0198	0	(30)	62	1	U	R.IVDIDPDGGNVSVTVR.V
<input checked="" type="checkbox"/>	<u>14</u>	859.9365	1717.8584	1717.8483	0.0102	0	(87)	0.00012	1	U	R.NGETTILSGVVADETGR.L
<input checked="" type="checkbox"/>	<u>15</u>	859.9412	1717.8678	1717.8483	0.0196	0	115	2.1e-07	1	U	R.NGETTILSGVVADETGR.L
<input checked="" type="checkbox"/>	<u>16</u>	967.0100	1932.0054	1931.9840	0.0215	0	47	1.3	1	U	K.SDLQNLQYVSVPLDEAK.Q
<input checked="" type="checkbox"/>	<u>18</u>	719.3674	2155.0804	2155.0698	0.0106	0	79	0.00072	1	U	R.LPFTDWAPRPDVEEGASLR.L
<input checked="" type="checkbox"/>	<u>19</u>	1486.2783	2970.5420	2970.4823	0.0598	0	86	9.5e-05	1	U	R.GVPQVNLSEFTTLDVLDDPVSVTDSAPR.L

Match to: [gi|292655492](#) Score: 940  
 replication protein A [Haloferax volcanii DS2]  
 Found in search of 1C\_MP2411\_14SEPT11.pk1  
 Nominal mass (M<sub>r</sub>): 45954; Calculated pI value: 4.20  
 NCBI BLAST search of [gi|292655492](#) against nr  
 Unformatted [sequence string](#) for pasting into other applications  
 Taxonomy: [Haloferax volcanii DS2](#)  
 Links to retrieve other entries containing this sequence from NCBI Entrez:  
[gi|291371312](#) from [Haloferax volcanii DS2](#)  
 Variable modifications: Carbamidomethyl (C),Oxidation (M)  
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Sequence Coverage: 40%  
 Matched peptides shown in **Bold Red**

```

1  MELDQHAEEEL  ASALGVDKEE  VKSDLQNLQ  YSVPLDEAKQ  SVRRKHGGGS
51  SGGSDGAPAT  KRIVDIDPDG  GNVSVTVRVL  TVGTRSIVYQ  GDEQTIREGE
101 LADESGVISY  TAWQDFGFEP  GDSVVIGNAG  VREWDGKPEL  NIGASSTVGV
151 ESETVETPYD  DRIGGEADLI  DLQAGDRGRT  VEVVRVLEVDS  RTIDGRNGET
201 TILSGVVADE  TGRLPFTDWA  PRPDVEEGAS  LRLSDVYVRE  FRGVPQVNLS
251 EFTTLDVLD  PVSVTDSAPR  LKIGEAVDAG  GMFDVELLGN  VLEVRDGSGL
301 IERCPKCSR  VQNGQCRQHG  EVDGEDDMRV  KAILDDGTGT  VTAILDRDLT
351 EEIYGGTMAD  AMEAAREAMD  KEVVADDIAD  TLVGREYRVR  GNLSVDEYGA
401 NLETDEFEET  DDDPAERAAA  LLTEVRA
```

1D\_MP2411 RPA1AP

gene="rpa1ap"  
 locus\_tag="HVO\_1337"



# SEARCHED FOR his tagged N-TERMINAL PEPTIDE, NO SIGN OF IT

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110915/FtocmaatL.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110915/FtocmaatL.dat)

1. [gi|292655491](#) Mass: 64829 Score: 578 Matches: 10(5) Sequences: 9(5)  
rpa-associated protein [Haloferax volcanii DS2]

☐ Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/> 5	433.2425	864.4704	864.4705	0.0000	0	27	2e+02	1	U	R.TYLEALR.E
<input checked="" type="checkbox"/> 7	517.2556	1032.4966	1032.4989	0.0022	0	35	27	1	U	R.AIDHYGTTR.T
<input checked="" type="checkbox"/> 8	571.8137	1141.6128	1141.6091	0.0037	0	73	0.0043	1	U	R.GVDSTLAAGVPR.A
<input checked="" type="checkbox"/> 10	608.8337	1215.6528	1215.6611	0.0083	0	56	0.22	1	U	R.WIVSAAEATLR.R
<input checked="" type="checkbox"/> 15	735.4022	1468.7898	1468.7885	0.0013	0	88	0.0001	1	U	R.EVAVQALELVAGDR.D
<input checked="" type="checkbox"/> 17	764.3859	1526.7572	1526.7439	0.0134	0	84	0.00029	1	U	R.IAVFDEALSMPYR.G + Oxidation (M)
<input checked="" type="checkbox"/> 19	709.7049	2126.0929	2126.1008	0.0079	0	(22)	3.2e+02	1	U	R.TFVAGVLTEVEHVNDVLR.G
<input checked="" type="checkbox"/> 20	1064.0586	2126.1026	2126.1008	0.0018	0	32	35	1	U	R.TFVAGVLTEVEHVNDVLR.G
<input checked="" type="checkbox"/> 21	1133.9921	2265.9696	2265.9549	0.0147	0	78	0.00074	1	U	R.VFAAEFDDASLSYSESDEER.A
<input checked="" type="checkbox"/> 23	1000.1385	2997.3937	2997.3840	0.0097	0	107	6.5e-07	1	U	R.TYEPDDADVVSVPESVNTVDADVR.D

Match to: [gi|292655491](#) Score: 578

rpa-associated protein [Haloferax volcanii DS2]

Found in search of 1D\_MP2411\_15SEPT11.pkl

Nominal mass (M<sub>r</sub>): 64829; Calculated pI value: 3.49

NCBI BLAST search of [gi|292655491](#) against nr

Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Haloferax volcanii DS2](#)

Links to retrieve other entries containing this sequence from NCBI Entrez:

[gi|291371037](#) from [Haloferax volcanii DS2](#)

Variable modifications: Carbamidomethyl (C),Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 21%

```

1 MSASPVVGTR EIAYRVFAAE FDDASLSYSE SDEERAPNYV VTPTGARLNR
51 TFVAGVLTEV EHVNDEVLRG RIADPTGAFV TYAGQYQPEP MAYLDAATPP
101 AFVSLAGKAR TYEPDDADV YSSVRPESVN TVDADVRDRW IVSAAEATLR
151 RIAVFDEALS MPYRGDDLTR ALEARGV DST LAAGVPR AID HYGTTTRTYLE
201 ALREVAVQAL ELVAGDRDQV DPLDVAPGDG GDAVLGPLPE LDLEPAASVD
251 IEVGGEADAGE VEADAGEFEN ESGSEPAATL DSEPEFDGGA ESEAVEPESE
301 SEPPIAEPDSE PESASDPEPV AAETTAAAES QPGASESEPA SEPSTTV DSE
351 PEPEPAAEPE SEPAAEPDLT SEPEPEPVT D SLDTESEPDF DDGALGDFED
401 DAPDAGATDA GSLDDSGSGT LGDFDDGFDD PDPEAGDSTD STDSASADAD
451 DSVDPDGMYE LDEDERAEIE SEFGTEFTSG ADVDPAGEAD IDVPDADDLT
501 EQLEDESAAA AESDPEPATA AAEPEPAVDS EPEPTPDAAS DGDEDEADAD
551 PDADESAEDI DLESVVVDAM DDLDDGDGAT RDEVVAAVVD EHGADPGAVE
601 DAIQEALLGG RCYEPQDGV L KAI

```

1E\_MP2411 RPA1

gene="rpa1"

locus\_tag="HVO\_1338"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110915/FtocmaEtO.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110915/FtocmaEtO.dat)

1. [gi|292655492](#) Mass: 45954 Score: 808 Matches: 15(7) Sequences: 12(7)  
replication protein A [Haloferax volcanii DS2]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
<input checked="" type="checkbox"/> 1	409.2268	816.4390	816.4341	0.0049	0	53	0.59	1	U	R.VLEVDSR.T
<input checked="" type="checkbox"/> 2	426.2339	850.4532	850.4549	0.0016	0	39	10	1	U	R.LSDVYVR.E
<input checked="" type="checkbox"/> 3	472.2832	942.5518	942.5498	0.0020	0	79	0.0011	1	U	R.AAALLTEVR.A

✓	4	693.7705	1385.5264	1386.5470	1.0206	0	(36)	17	1	U	R.QHGEVDGEDDMR.V
✓	6	702.2792	1402.5438	1402.5419	0.0019	0	71	0.0055	1	U	R.QHGEVDGEDDMR.V + Oxidation (M)
✓	7	704.8595	1407.7044	1407.6994	0.0050	0	73	0.0033	1	U	R.SIVYQGDEQTIR.E
✓	8	705.3808	1408.7470	1407.6994	1.0476	0	(18)	9.9e+02	1	U	R.SIVYQGDEQTIR.E
	10	736.8851	1471.7556	1471.7518	0.0038	0	3	3.9e+04	10	U	K.EVVADDIADTLVGR.E
✓	12	815.9409	1629.8672	1629.8574	0.0099	0	123	3e-08	1	U	K.AILDDGTGTVTAILDR.D
✓	13	828.4398	1654.8650	1654.8526	0.0124	0	40	5.8	1	U	R.IVDIDPDGGNVSVTVR.V
✓	14	860.4330	1718.8514	1717.8483	1.0032	0	111	4.9e-07	1	U	R.NGETTILSGVVADETGR.L
	15	967.0059	1931.9972	1931.9840	0.0133	0	(3)	3e+04	7	U	K.SDLQNLQYSVPLDEAK.Q
✓	16	967.0076	1932.0006	1931.9840	0.0167	0	81	0.00046	1	U	K.SDLQNLQYSVPLDEAK.Q
✓	17	1038.4677	2074.9208	2074.8823	0.0385	0	22	3.1e+02	1	U	R.DLTEEYGGTMADAMEAAR.E + 2 Oxidation (M)
✓	19	1486.2662	2970.5178	2970.4823	0.0356	0	115	1.1e-07	1	U	R.GVPQVNLSEFTTLDVLDDFVSVTDSAPR.L

Match to: [gi|292655492](#) Score: 808  
**replication protein A [Haloferax volcanii DS2]**  
Found in search of 1E\_MP2411\_15SEPT11.pkl  
Nominal mass (M<sub>r</sub>): **45954**; Calculated pI value: **4.20**  
NCBI BLAST search of [gi|292655492](#) against nr  
Unformatted [sequence string](#) for pasting into other applications  
Taxonomy: [Haloferax volcanii DS2](#)  
Links to retrieve other entries containing this sequence from NCBI Entrez:  
[gi|291371312](#) from [Haloferax volcanii DS2](#)  
Variable modifications: Carbamidomethyl (C),Oxidation (M)  
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
Sequence Coverage: **40%**  
Matched peptides shown in **Bold Red**

1 MELDQHAEEL ASALGVDKEE VK**SDLQNLQ YSVPLDEAKQ** SVRRKHGGGS  
51 SGGSDGAPAT KR**IVDIDPDG GNVSVTVRVL** TVGTR**SIVYQ GDEQTIREGE**  
101 LADESGVISY TAWQDFGFEP GDSVIGNAG VREWDGKPEL NIGASSTVG  
151 ESETVETPYD DRIGGEADLI DLQAGDRGRT VEV**RVLEVDS RTIDGRNGET**  
201 **TILSGVVADE TGR**LPFTDWA PRPDVEEGAS LR**LSDVYVRE FRGVPQVNLS**  
251 **EFTTLDVLDD PVS**VTDSAPR LKIGEAVDAG GMFDVLLGN VLEVRDGSGL  
301 IERCPKCSR VQNGQCR**QHG EVDGEDDMRV KAILDDGTGT VTAILDRDLT**  
351 **EEIYGGTMAD AMEAAREAMD KEVVADDIAD TLVGR**EYRVR GNLSVDEYGA  
401 NLETDEFEET DDDPAER**AAA LLTEVRA**

## Rpa3AP and RPA3AP MP\_2511

### 1H\_MP2511 RPA3AP

gene="rpa3ap"

locus\_tag="HVO\_0291"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110926/FtociNeah.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110926/FtociNeah.dat)

1. [gi|292654471](#) Mass: 21979 Score: 446 Matches: 9 (4) Sequences: 8 (4)  
rpa-associated protein [Haloferax volcanii DS2]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
✓	4	475.2297	948.4448	948.4553	-0.0104	0	33	49	1	U R.VWTETADK.T
✓	8	706.2849	1410.5552	1410.5688	-0.0135	0	60	0.059	1	U K.EDVGEDNEYWR.G
✓	9	743.3309	1484.6472	1484.6419	0.0053	0	97	1.5e-05	1	U R.IAAFDEEGDEYAR.M
✓	10	743.3350	1484.6554	1484.6419	0.0135	0	(18)	1.1e+03	1	U R.IAAFDEEGDEYAR.M
✓	12	785.9380	1569.8614	1569.8515	0.0100	0	66	0.019	1	U R.DLDAPAYVAVVGKPR.T
	13	531.9094	1592.7064	1592.7107	-0.0043	1	18	1e+03	2	U R.EHYDLDPPEYKR.A
✓	14	800.9406	1599.8666	1599.8620	0.0046	0	68	0.012	1	U R.APVYLLPTGESANR.V
✓	16	818.8895	1635.7644	1635.7569	0.0075	0	69	0.0089	1	U R.VFAQEFNDAGYTFK.E
✓	18	851.4364	1700.8582	1700.8468	0.0114	0	36	17	1	U R.AAIAALESLEQADELSA.-

Match to: [gi|292654471](#) Score: 446  
**rpa-associated protein [Haloferax volcanii DS2]**  
Found in search of 1H\_MP2511\_26SEPT11.pkl  
Nominal mass (M<sub>r</sub>): **21979**; Calculated pI value: **4.30**

NCBI BLAST search of [gi|292654471](#) against nr  
 Unformatted [sequence string](#) for pasting into other applications  
 Taxonomy: [Haloferax volcanii DS2](#)  
 Links to retrieve other entries containing this sequence from NCBI Entrez:  
[gi|291370388](#) from [Haloferax volcanii DS2](#)  
 Variable modifications: Carbamidomethyl (C),Oxidation (M)  
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Sequence Coverage: 53%  
 Matched peptides shown in **Bold Red**

1 MSSNEIPTRE VARR**VFAQEF** NDAGYTFKES DDERAPVYLL LPTGESANRV  
 51 FLVGTLTEKE **DVGEDNEYWR** GRIVDPTGTF FVYAGQYQPE AASALR**DLDA**  
 101 **PAYVAVVGKP** RTYETDDGSI NVSVRPESIT EVDAATRDRW **VTETADK**TLD  
 151 **RIAADFDEEGD** EYARMAREHY **DLDPPEYKRA** AIAALESLEQ **ADELSA**

# 1D\_MP2511 RPA3 and Thioredoxin reductase

gene="rpa3"

locus\_tag="HVO\_0292"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110927/FtoclFSeR.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110927/FtoclFSeR.dat)

1.	<a href="#">gi 292654472</a>	Mass: 34562	Score: 293	Matches: 6(2)	Sequences: 5(2)
replication protein A [Haloferax volcanii DS2]					
	Query	Observed	Mr(expt)	Mr(calc)	Delta Miss Score Expect Rank Unique Peptide
<input checked="" type="checkbox"/>	7	748.8972	1495.7798	1495.8246	-0.0447 1 (15) 2.4e+03 7 U R.LREPADAELLIK.A
<input checked="" type="checkbox"/>	8	748.9021	1495.7896	1495.8246	-0.0349 1 59 0.081 1 U R.LREPADAELLIK.A
<input checked="" type="checkbox"/>	10	776.3713	1550.7280	1550.7576	-0.0296 0 85 0.0002 1 U R.SVVNSYLDEAGIER.D
<input checked="" type="checkbox"/>	11	802.3829	1602.7512	1602.7849	-0.0337 0 88 0.00011 1 U R.SESIAQVGLLGDDSGR.M
<input checked="" type="checkbox"/>	13	854.3883	1706.7620	1706.7920	-0.0300 0 7 1.2e+04 3 U R.EMTEELTGIELDEAK.Q
<input checked="" type="checkbox"/>	16	906.9296	1811.8446	1811.8829	-0.0383 0 56 0.15 1 U K.FVSFTTSELPELEEGK.S

Match to: [gi|292654472](#) Score: 293  
**replication protein A [Haloferax volcanii DS2]**  
 Found in search of 1D\_MP2511\_26SEPT11.pk1  
 Nominal mass (M<sub>r</sub>): **34562**; Calculated pI value: **4.20**  
 NCBI BLAST search of [gi|292654472](#) against nr  
 Unformatted [sequence string](#) for pasting into other applications  
 Taxonomy: [Haloferax volcanii DS2](#)  
 Links to retrieve other entries containing this sequence from NCBI Entrez:  
[gi|291371038](#) from [Haloferax volcanii DS2](#)  
 Variable modifications: Carbamidomethyl (C),Oxidation (M)  
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Sequence Coverage: 23%  
 Matched peptides shown in **Bold Red**

1 MTDLRTHAAE IADQFSDHLD VSADEVEERL ESLVTEYRVP VDEARR**SVVN**  
 51 **SYLDEAGIER** DELAGGSGGN EQTLNDIDE DEQWVDVRK VVELWEPR**SE**  
 101 **SIAQVGLLGD** DSGRMK**FVSF** **TTSELPELEE** GKSVALGNV TDEYQGNFSV  
 151 KLNRTTSITE LDEEIEVGDD STSVEGALVD SQSGSGLIKR CPEEGCTRVL  
 201 QNGRCSEHGS VEGFDLRIK AVVDDGDEVH EVIFN**EMTE** **ELTGIELDEA**  
 251 **KQAMDALDT** TIVEEEMRGD LVGYYRVVTG PTLGRYVLN EVER**LREPAD**  
 301 **AEELLIKARS** M

2.	<a href="#">gi 292655899</a>	Mass: 36334	Score: 91	Matches: 3(0)	Sequences: 3(0)
thioredoxin reductase [Haloferax volcanii DS2]					
	Query	Observed	Mr(expt)	Mr(calc)	Delta Miss Score Expect Rank Unique Peptide
<input checked="" type="checkbox"/>	4	509.2795	1016.5444	1016.5502	-0.0058 0 32 62 1 U R.LVETIGTER.V
<input checked="" type="checkbox"/>	9	767.3926	1532.7706	1532.8172	-0.0465 0 36 18 1 U R.NAQVGIAGVGQGHVAR.C
<input checked="" type="checkbox"/>	18	923.4493	1844.8840	1844.9156	-0.0315 0 23 3.1e+02 1 U R.VLDALPDEALADYLDGR.E

Match to: [gi|292655899](#) Score: 91  
**thioredoxin reductase [Haloferax volcanii DS2]**  
 Found in search of 1D\_MP2511\_26SEPT11.pk1  
 Nominal mass (M<sub>r</sub>): **36334**; Calculated pI value: **4.30**  
 NCBI BLAST search of [gi|292655899](#) against nr  
 Unformatted [sequence string](#) for pasting into other applications  
 Taxonomy: [Haloferax volcanii DS2](#)  
 Links to retrieve other entries containing this sequence from NCBI Entrez:  
[gi|291372954](#) from [Haloferax volcanii DS2](#)  
 Variable modifications: Carbamidomethyl (C),Oxidation (M)  
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Sequence Coverage: 12%  
 Matched peptides shown in **Bold Red**

1 MTDQGTLEDS ADVVVVGSGP SGCAAVFTA RYGLDVTTFD RGNAALRRCA  
51 YLENYLGFP GIGVDTFTAL MHEHLAEVGA DYVADMVSV ERPGDESEVA  
101 AASEGESDAF DADARRFVVR TQEGRRVEAD HVVAAAWYDG DYLRGLDDDD  
151 AMFEEHDDHG EVHEQFDSY ADDDGRTPVD GLYVAAPSGD **RNAQVGIAVG**  
201 **QGGHVAR**CLI EDRRRERGF GDLAHYDWL RADTEFHGEW GDRDRWREWF  
251 DGEVPDDADL DDERLAELRE SYIDRAFDTR RTDHEVEAAE PRGIRRLVET  
301 **IGTERVLDAL PDEALADYLD GRELGGVEQ**

#### 1C\_MP2511 Deoxyhypusine synthase

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110927/FtocIfSnL.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110927/FtocIfSnL.dat)

1.	gi 292656425		Mass: 38616	Score: 124	Matches: 6(0)	Sequences: 5(0)					
	deoxyhypusine synthase [Haloferax volcanii DS2]										
	Query	Observed	Mr(expt)	Mr(calc)	Delta	MissScore	Expect	Rank	Unique	Peptide	
<input checked="" type="checkbox"/>	6	473.5289	1417.5649	1417.6335	-0.0686	0	37	14	1	U	R.AHPDPSSSPSHGSR.E
<input checked="" type="checkbox"/>	8	548.2188	1641.6346	1641.7230	-0.0885	1	29	65	1	U	R.EHDETLRDEEVDR.I
<input checked="" type="checkbox"/>	10	1020.5397	2039.0648	2039.1449	-0.0800	0	(7)	1.2e+04	1	U	R.NTTVLGDATIMLPLLVAAR.E
<input checked="" type="checkbox"/>	11	1028.5319	2055.0492	2055.1398	-0.0906	0	29	72	1	U	R.NTTVLGDATIMLPLLVAAR.E + Oxidation (M)
	14	1111.5000	2220.9854	2221.0573	-0.0718	0	2	3.5e+04	6	U	R.SEVFPALAEEGTVSIADLCR.E + Carbamidomethyl (C)
<input checked="" type="checkbox"/>	15	825.7190	2474.1352	2474.2401	-0.1050	0	27	96	1	U	R.DGHVDALVTTGANLTHDAIEAIGGK.H

Match to: **gi|292656425** Score: 124  
**deoxyhypusine synthase [Haloferax volcanii DS2]**  
Found in search of 1C\_MP2511\_26SEPT11.pkl  
Nominal mass ( $M_r$ ): **38616**; Calculated pI value: **4.45**  
NCBI BLAST search of [gi|292656425](#) against nr  
Unformatted [sequence string](#) for pasting into other applications  
Taxonomy: [Haloferax volcanii DS2](#)  
Links to retrieve other entries containing this sequence from NCBI Entrez:  
[gi|291371516](#) from [Haloferax volcanii DS2](#)  
Variable modifications: Carbamidomethyl (C), Oxidation (M)  
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
Sequence Coverage: 25%  
Matched peptides shown in **Bold Red**

1 MSDHDDDDGYE MPDREEFQHD PLGHAEVRRG MTVGELVEQY GHAGIGAADV  
51 HEADIIYAEM LADDCTVFM SLAGAMVPTG MRKVVSILIR **DGHVDALVTT**  
101 **GANLTHDAIE AIGGKHHHGR AHPDPSSSPS HGSREHGAEH HGEKTEREHD**  
151 **ETLRDEEVDR** IYNVYLPQEH FALFESHLRS **EVFPALAEETG TVSIADLCRE**  
201 LGRANSEVND REGIDEDAGV AAAAYEADVP IYCPAVQDSV LGLQAWMYSQ  
251 TSSFSLDALA DMTPLTDLAY DADTAGCLLV GGGVPKNFTL QTMLVTPGAY  
301 DYGVQITMDP AATGGLSGAT LDEARSWGKL EKDAR**NTTVL** **GDATIMLPLL**  
351 **VAAARERIE**

#### 1G\_MP2511 RPA3AP

gene="rpa3ap"

locus\_tag="HVO\_0291"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110927/FtocIfSec.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110927/FtocIfSec.dat)

1.

gi|292654471

Mass: 21979

Score: 435

Matches: 10(4)

Sequences: 7(4)

rpa-associated protein [Haloferax volcanii DS2]

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Unique	Peptide	
<input checked="" type="checkbox"/>	<u>3</u>	475.2198	948.4250	948.4553	-0.0302	0	32	54	1	U	R.WVTETADK.T
<input checked="" type="checkbox"/>	<u>6</u>	743.2962	1484.5778	1484.6419	-0.0641	0	84	0.00025	1	U	R.IAAFDEEGDEYAR.M
<input checked="" type="checkbox"/>	<u>7</u>	743.2999	1484.5852	1484.6419	-0.0567	0	(52)	0.44	1	U	R.IAAFDEEGDEYAR.M
<input checked="" type="checkbox"/>	<u>11</u>	531.8835	1592.6287	1592.7107	-0.0820	1	48	0.99	1	U	R.EHYDLDPPEYKR.A
<input checked="" type="checkbox"/>	<u>12</u>	800.9042	1599.7938	1599.8620	-0.0682	0	(56)	0.19	1	U	R.APVYLLPTGESANR.V
<input checked="" type="checkbox"/>	<u>13</u>	800.9050	1599.7954	1599.8620	-0.0666	0	70	0.0071	1	U	R.APVYLLPTGESANR.V
<input checked="" type="checkbox"/>	<u>14</u>	818.8504	1635.6862	1635.7569	-0.0707	0	(48)	0.95	1	U	R.VFAQEFNDAGYTFK.E
<input checked="" type="checkbox"/>	<u>15</u>	818.8519	1635.6892	1635.7569	-0.0677	0	82	0.00037	1	U	R.VFAQEFNDAGYTFK.E
<input checked="" type="checkbox"/>	<u>16</u>	851.4065	1700.7784	1700.8468	-0.0684	0	55	0.21	1	U	R.AAIAALESLEQAELSA.-
<input checked="" type="checkbox"/>	<u>21</u>	942.4132	2824.2178	2824.3363	-0.1185	0	63	0.016	1	U	R.TYETDDGSINVSVPESITEVDAATR.D

Match to: **gi|292654471** Score: 435  
**rpa-associated protein [Haloferax volcanii DS2]**  
Found in search of 1G\_MP2511\_26SEPT11.pkl

Nominal mass ( $M_r$ ): **21979**; Calculated pI value: **4.30**  
 NCBI BLAST search of [gi|292654471](#) against nr  
 Unformatted [sequence string](#) for pasting into other applications  
 Taxonomy: [Haloferax volcanii DS2](#)  
 Links to retrieve other entries containing this sequence from NCBI  
 Entrez:  
[gi|291370388](#) from [Haloferax volcanii DS2](#)  
 Variable modifications: Carbamidomethyl (C),Oxidation (M)  
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Sequence Coverage: **53%**  
 Matched peptides shown in **Bold Red**

1 MSSNEIPTRE VARR**VFAQEF NDAGYTFKES** DDER**APVYLL LPTGESANRV**  
 51 FLVGTLT**KE DVGEDNEYWR** GRIVDPTGTF FVYAGQYQPE AASALRDLDA  
 101 PAYVAVVGKP **RTYETDDGSI NVSVRPESIT EVDAATRDRW VTETADKTLT**  
 151 R**IAAFDEEGD EYARMAREHY DLDPEEYKRA AIAALESLEQ ADELSA**

#### 1B\_MP2511 RPA3AP

gene="rpa3ap"

locus\_tag="HVO\_0291"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=./data/20110927/FtocIfSnt.dat](http://www.matrixscience.com/cgi/master_results.pl?file=./data/20110927/FtocIfSnt.dat)

1.	gi 292654471	Mass: 21979	Score: 484	Matches: 10(4)	Sequences: 8(4)
	rpa-associated protein [Haloferax volcanii DS2]				
Query	Observed	Mr(expt)	Mr(calc)	Delta Miss	Score Expect Rank Unique Peptide
<input checked="" type="checkbox"/> 5	475.2141	948.4136	948.4553	-0.0416 0	39 11 1 U R.WVTETADK.T
<input checked="" type="checkbox"/> 8	706.2626	1410.5106	1410.5688	-0.0581 0	95 2.1e-05 1 U K.EDVGEDNEYWR.G
<input checked="" type="checkbox"/> 9	743.2981	1484.5816	1484.6419	-0.0603 0	88 9.6e-05 1 U R.IAAFDEEGDEYAR.M
<input checked="" type="checkbox"/> 10	743.3043	1484.5940	1484.6419	-0.0479 0	(30) 59 1 U R.IAAFDEEGDEYAR.M
<input checked="" type="checkbox"/> 12	531.8860	1592.6362	1592.7107	-0.0745 1	26 1.7e+02 1 U R.EHYDLDPEEYKR.A
<input checked="" type="checkbox"/> 13	800.9046	1599.7946	1599.8620	-0.0674 0	(28) 1e+02 1 U R.APVYLLPTGESANR.V
<input checked="" type="checkbox"/> 14	800.9059	1599.7972	1599.8620	-0.0648 0	68 0.011 1 U R.APVYLLPTGESANR.V
<input checked="" type="checkbox"/> 15	818.8519	1635.6892	1635.7569	-0.0677 0	88 9.2e-05 1 U R.VFAQEFNDAGYTFK.E
<input checked="" type="checkbox"/> 16	851.3970	1700.7794	1700.8468	-0.0674 0	40 5.8 1 U R.AAIAALESLEQADELSA.-
<input checked="" type="checkbox"/> 19	942.4132	2824.2178	2824.3363	-0.1185 0	40 4 1 U R.TYETDDGSINVSVPESITEVDAATR.D

Match to: [gi|292654471](#) Score: 484

rpa-associated protein [Haloferax volcanii DS2]

Found in search of 1B\_MP2511\_26SEPT11.pk1

Nominal mass ( $M_r$ ): **21979**; Calculated pI value: **4.30**

NCBI BLAST search of [gi|292654471](#) against nr

Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Haloferax volcanii DS2](#)

Links to retrieve other entries containing this sequence from NCBI Entrez:

[gi|291370388](#) from [Haloferax volcanii DS2](#)

Variable modifications: Carbamidomethyl (C),Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: **59%**

Matched peptides shown in **Bold Red**

1 MSSNEIPTRE VARR**VFAQEF NDAGYTFKES** DDER**APVYLL LPTGESANRV**  
 51 FLVGTLT**KE DVGEDNEYWR** GRIVDPTGTF FVYAGQYQPE AASALRDLDA  
 101 PAYVAVVGKP **RTYETDDGSI NVSVRPESIT EVDAATRDRW VTETADKTLT**  
 151 R**IAAFDEEGD EYARMAREHY DLDPEEYKRA AIAALESLEQ ADELSA**

#### 1A\_MP2511 RPA3

gene="rpa3"

locus\_tag="HVO\_0292"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=./data/20110927/FtocIfSnR.dat](http://www.matrixscience.com/cgi/master_results.pl?file=./data/20110927/FtocIfSnR.dat)

1.	gi 292654472	Mass: 34562	Score: 796	Matches: 13(6)	Sequences: 10(6)
	replication protein A [Haloferax volcanii DS2]				
Query	Observed	Mr(expt)	Mr(calc)	Delta Miss	Score Expect Rank Unique Peptide
<input checked="" type="checkbox"/> 2	400.7177	799.4208	799.4552	-0.0344 0	55 0.26 1 U R.VTGPTLGR.Y
<input checked="" type="checkbox"/> 4	546.7614	1091.5082	1091.5611	-0.0528 0	53 0.46 1 U R.YVLANEVER.L
<input checked="" type="checkbox"/> 5	555.2692	1108.5238	1108.5764	-0.0526 0	39 8.8 1 U R.LESLVTEYR.V
<input checked="" type="checkbox"/> 7	748.8840	1495.7534	1495.8246	-0.0711 1	(48) 1.1 1 U R.LREPADAEEELLIK.A

✓	8	748.8843	1495.7540	1495.8246	-0.0705	1	84	0.00029	1	U	R.LREPADAELLIK.A
✓	9	776.3477	1550.6808	1550.7576	-0.0768	0	60	0.074	1	U	R.SVNSYLDDEAGIER.D
✓	10	802.3569	1602.6992	1602.7849	-0.0857	0	(39)	9.7	1	U	R.SESIAQVGLLGDDSGR.M
✓	11	802.3614	1602.7082	1602.7849	-0.0767	0	117	1.3e-07	1	U	R.SESIAQVGLLGDDSGR.M
✓	14	854.3666	1706.7186	1706.7920	-0.0734	0	84	0.00027	1	U	R.EMTEELTGIELDEAK.Q
✓	15	906.9073	1811.8000	1811.8829	-0.0829	0	80	0.00057	1	U	K.FVSFTTSELPELEEGK.S
✓	17	991.9012	1981.7878	1981.8795	-0.0916	0	(4)	1.9e+04	7	U	K.QMAMDALDTTIVEEEMR.G
✓	18	991.9046	1981.7946	1981.8795	-0.0848	0	117	9.1e-08	1	U	K.QMAMDALDTTIVEEEMR.G
✓	19	1045.9618	2089.9090	2089.9957	-0.0866	0	110	5.2e-07	1	U	K.SYALGNVVTDEYQGNFSVK.L

Match to: [gi|292654472](#) Score: 796  
**replication protein A [Haloferax volcanii DS2]**  
Found in search of 1A\_MP2511\_26SEPT11.pkl  
Nominal mass (M<sub>r</sub>): **34562**; Calculated pI value: **4.20**  
NCBI BLAST search of [gi|292654472](#) against nr  
Unformatted [sequence string](#) for pasting into other applications  
Taxonomy: [Haloferax volcanii DS2](#)  
Links to retrieve other entries containing this sequence from NCBI Entrez:  
[gi|291371038](#) from [Haloferax volcanii DS2](#)  
Variable modifications: Carbamidomethyl (C),Oxidation (M)  
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
Sequence Coverage: **43%**  
Matched peptides shown in **Bold Red**  
1 MTDLRTHAAE IADQFSDHLD VSADEVEERL **ESLVTEYRVP** VDEARRSVVN  
51 **SYLDEAGIER** DELAGGSGGN EQTLNDIDE DEQWVDVRK VVELWEPRSE  
101 **SIAQVGLLGD** DSGRMKFVSF **TTSELPELEE** GKSALGNVV **TDEYQGNFSV**  
151 **KLNRRTTSITE** LDEEIEVGDD STSVEGALVD SQSGSGLIKR CPEEGCTRVL  
201 QNGRCSEHGS VEGEFDLRIK AVVDDGDEVH EVIFNREMT **ELTGIELDEA**  
251 **QAMAMDALDT** **TIVEEEMRGD** LVGYYYRVTG **PTLGRYVLAN** EVERLREPAD  
301 **AEELLIKARS** M

#### 1E\_MP2511 RPA3 and Thioredoxin reductase

gene="rpa3"

locus\_tag="HVO\_0292"

[http://www.matrixscience.com/cgi/master\\_results.pl?file=../data/20110926/FtocInaEE.dat](http://www.matrixscience.com/cgi/master_results.pl?file=../data/20110926/FtocInaEE.dat)

Match to: [gi|292654472](#) Score: 735  
**replication protein A [Haloferax volcanii DS2]**  
Found in search of 1E\_MP2511\_26SEPT11.pkl  
Sequence Coverage: **48%**  
Matched peptides shown in **Bold Red**

1 MTDLRTHAAE IADQFSDHLD VSADEVEERL **ESLVTEYRVP** VDEARRSVVN  
51 **SYLDEAGIER** DELAGGSGGN EQTLNDIDE DEQWVDVRK VVELWEPRSE  
101 **SIAQVGLLGD** DSGRMKFVSF **TTSELPELEE** GKSALGNVV **TDEYQGNFSV**  
151 **KLNRRTTSITE** LDEEIEVGDD STSVEGALVD SQSGSGLIKR CPEEGCTRVL  
201 QNGRCSEHGS VEGEFDLRIK AVVDDGDEVH EVIFNREMT **ELTGIELDEA**  
251 **QAMAMDALDT** **TIVEEEMRGD** LVGYYYRVTG **PTLGRYVLAN** EVERLREPAD  
301 **AEELLIKARS** M

Match to: [gi|292655899](#) Score: 140  
**thioredoxin reductase [Haloferax volcanii DS2]**  
Found in search of 1E\_MP2511\_26SEPT11.pkl  
Sequence Coverage: **10%**  
Matched peptides shown in **Bold Red**

1 MTDQGTEDLS ADVVVVGSGP SGCAAAVFTA RYGLDTPVFD RGNAALRRCA  
51 YLENYLGFP A GIGVDTFTAL MHEHLAEVGA DYVADMVSV ERPGDESEVA  
101 AASEGESDAF DADARRFVVR TQEGRRVEAD HVVAAAWYDG DYLRGLDDDD  
151 AMFEEHDHGG EVHEQFDSY ADDDGRTPVD GLYVAAPSGD **RNAQVGI AVG**  
201 **OGGHVARCLI** EDRRRERGF GDLAAHYDWL RADTEFHGEW GDRDRWREWF  
251 DGEVPDDADL DDERLAELRE SYIDRAFDTR RTDHEVEAAE PRGIRRLVET  
301 IGTERVLDAL **PDEALADYLD** **GRELGGVEQ**

## Note in proof

The paper, 'Identification of essential and non-essential single-stranded DNA-binding proteins in a model archaeal organism', Skowyra A, Macneill SA, Nucleic Acids Research, pages 1 -14, was published on the 5<sup>th</sup> October 2011.

This paper identifies that RPA1 and RPA3 have overlapping functions but that RPA2 is essential. This study also investigates the roles of the N and C-terminus and the individual oligosaccharide-binding fold of RPA2. This is achieved through examining the growth and sensitivity to DNA damage of partial  $\Delta rpa2$  deletion mutants. Interestingly an increased level of expression of *rpa3* is shown to rescue loss of RPA2 function and an elevated level of RPA2 is thought to increase DNA damage resistance.